

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: September 12, 2003, 20:29:05 ; Search time 1868.44 Seconds

(without alignments)  
17056.227 Million cell updates/sec

Title: US-09-847-232a-26

Perfect score: 1 ttcctagtagtaagcaaac.....ttcattagtcctacaacac 779

Sequence: 1 ttcctagtagtaagcaaac.....ttcattagtcctacaacac 779

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenBank.\*  
1: gb\_ba.\*  
2: gb\_hg.\*  
3: gb\_in.\*  
4: gb\_lm.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_om.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_to.\*  
27: em\_sts.\*  
28: em\_un.\*  
29: em\_vi.\*  
30: em\_hg\_hum.\*  
31: em\_hg\_inv.\*  
32: em\_hg\_other.\*  
33: em\_hg\_mus.\*  
34: em\_hg\_pln.\*  
35: em\_hg\_rtd.\*  
36: em\_hg\_mam.\*  
37: em\_hg\_vrt.\*  
38: em\_sy.\*  
39: em\_hggo\_hum.\*  
40: em\_hggo\_mus.\*  
41: em\_hggo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	779	100.0	779	6	AX298113
2	779	100.0	2774	8	RICGTLA
3	779	100.0	3435	8	OSGLTGE
4	763.4	98.0	839	6	120872
5	763.4	98.0	1240	6	120873
6	761.8	97.8	5494	8	RICGLUTE
7	761.8	97.8	144593	8	AC021891
8	761.8	97.8	324367	8	AE017091
9	729.4	93.6	910	8	RICGLTB
10	729.4	93.6	912	8	RICGLTF
11	717.4	92.1	910	8	RICGLTG
12	392.2	50.3	2867	8	RICGTA
13	303.6	39.0	1111	8	RICGLTC
14	289.4	37.2	481	8	RICGLTD
15	289.2	37.1	1119	8	RICGLTA
16	287.8	36.9	170021	8	AE003256
17	287.8	36.9	176261	8	AE003274
18	282.6	36.3	1116	8	RICGLTE
19	164.8	21.2	716	8	OSGLTGE
20	132	16.9	449	8	RICCONS
21	86.6	11.1	2439	8	RICGLTA
22	85.4	11.0	126824	2	AC133398
23	84.2	10.8	2936	8	OSGLTGE
24	84.2	10.8	106267	2	AC137999
25	54.2	7.0	4169	2	AC014614
26	50.8	6.5	2000	6	AX655393
27	48	6.2	61020	6	AX251546
28	47.6	6.1	254436	3	AE014827
29	47.2	6.1	252420	3	AE014841
30	46.2	5.9	204903	2	AL713986
31	45.8	5.9	157057	2	EX005474
32	45.8	5.9	189541	2	AC101785
33	45.8	5.9	233321	10	AL732546
34	45.6	5.9	1141	6	AX083744
35	45.2	5.8	250743	3	AE014836
36	45	5.8	55457	2	AC137139
37	45	5.8	121958	9	AL583838
38	45	5.8	128294	9	AC073248
39	44.8	5.8	34980	6	AX344551
40	44.6	5.7	78156	2	AC025101
41	44.6	5.7	180097	9	AC123978
42	44.6	5.7	186165	9	AC112187
43	44.4	5.7	144260	9	AC006998
44	44.4	5.7	184333	2	AC13898
45	44.4	5.7	188076	2	AC023629

## ALIGNMENTS

RESULT 1	AX298113	779 bp	DNA	linear	PAT 26-NOV-2001
LOCUS	AX298113				
DEFINITION	Sequence 26 from Patent WO0183792.				
ACCESSION	AX298113				
VERSION	AX298113.1				
KEYWORDS	GI:17128189				
SOURCE					
ORGANISM	Oryza sativa				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Euphorbiaceae; Oryzaceae; Oryza.				
REFERENCE	Huang, N., Huang, Y.S., Yang, D. and Schmidt, R.J.				
AUTHORS	1				
TITLE	Plant transcription factors and enhanced gene expression				

JOURNAL Patent: WO 0183792-A 26 08-NOV-2001;  
Applied Phytologics, Inc. (US)  
FEATURES Location/Qualifiers  
source 1. .779  
/organism="Oryza sativa"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:4530"  
BASE COUNT 285 a 159 c 123 g 212 t  
ORIGIN  
Query Match 100.0%; Score 779; DB 6; Length 779;  
Best Local Similarity 100.0%; Pred. No. 8 4e-166;  
Matches 779; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 TTCTGTAGTACAGCAAACTAAAGTAAGAGAGTGTGTGTAGAAAAAGAAC 60  
DB 1 TTCTGTAGTACAGCAAACTAAAGTAAGAGAGTGTGTGTAGAAAAAGAAC 60  
OY 61 AATATCATGATAGTGTGAGACATTAATGGACCAAGAAATAAAGAACTTTGTATGA 120  
DB 61 AATATCATGATAGTGTGAGACATTAATGGACCAAGAAATAAAGAACTTTGTATGA 120  
OY 121 GTGCTGATCTCGATGAGCCTCAAAAGTTCTCTCACCCGGATAGAAACCTTAAGCA 180  
DB 121 GTGCTGATCTCGATGAGCCTCAAAAGTTCTCTCACCCGGATAGAAACCTTAAGCA 180  
OY 181 ATGTGCAAGTTGATCTCCACTGACATTAAGCAAAATAGATATCATGATGACATA 240  
DB 181 ATGTGCAAGTTGATCTCCACTGACATTAAGCAAAATAGATATCATGATGACATA 240  
OY 241 GCAACATGACATCATATCAGCCCTGCTCAACTATGATCTCACTCATCAATCAATAG 300  
DB 241 GCAACATGACATCATATCAGCCCTGCTCAACTATGATCTCACTCATCAATCAATAG 300  
OY 301 TATCTTCAGCTAAATGTTAGAACATAAACCATAACGTTGATGAGATTAAGCGT 360  
DB 301 TATCTTCAGCTAAATGTTAGAACATAAACCATAACGTTGATGAGATTAAGCGT 360  
OY 361 GACACATGACAAATTCAGACTCAAGCAATTAACCAAAATGATGTGATCAATTAAC 420  
DB 361 GACACATGACAAATTCAGACTCAAGCAATTAACCAAAATGATGTGATCAATTAAC 420  
OY 421 CAGAGCTATATGTCATTAATGCAAAAAGAGAGCTTATTAAGACAAGCATGATCACA 480  
DB 421 CAGAGCTATATGTCATTAATGCAAAAAGAGAGCTTATTAAGACAAGCATGATCACA 480  
OY 481 AAATTCACCTGCTTCGTCGCAAAAAGAGAGGCTTACATTAATCCATGTCATATGC 540  
DB 481 AAATTCACCTGCTTCGTCGCAAAAAGAGAGGCTTACATTAATCCATGTCATATGC 540  
OY 541 AAAAGAAAGAGAAAGAAACAACAATGCTGCTCAATTAATCAATCTGATGCTCAAT 600  
DB 541 AAAAGAAAGAGAAAGAAACAACAATGCTGCTCAATTAATCAATCTGATGCTCAAT 600  
OY 601 CATTATTCATCACTCTTCGATGACACACTTCATATATCATTAAGAGTCACTTACGCT 660  
DB 601 CATTATTCATCACTCTTCGATGACACACTTCATATATCATTAAGAGTCACTTACGCT 660  
OY 661 GGAGATTAAACAACTCTATCTTAACATTTAGAGGAGGCTTTATCTCAGTAAAG 720  
DB 661 GGAGATTAAACAACTCTATCTTAACATTTAGAGGAGGCTTTATCTCAGTAAAG 720  
OY 721 CAGCATGATTCATGTTCTGCACAAAAGCATTCATTAAGTCACTAGCTAGCAAC 779  
DB 721 CAGCATGATTCATGTTCTGCACAAAAGCATTCATTAAGTCACTAGCTAGCAAC 779  
RESULT 2  
LOCUS RIGG1A 2774 bp DNA linear PLN 06-OCT-1994  
DEFINITION Oryza sativa glutelin 1 (Gt1) gene, complete cds.  
ACCESSION M28156  
VERSION M28156.1 GI:341838

KEYWORDS glutelin.  
SOURCE Oryza sativa (japonica cultivar-group)  
ORGANISM Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehharoidae; Oryzae; Oryza.  
REFERENCE 1 (bases 1 to 2774)  
Okita,T.W., Hwang,Y.S., Hnilo,J., Kim,W.T., Aryan,A.P., Larson,R.  
and Krishnan,H.B.  
TITLE Structure and expression of the rice glutelin multigene family  
JOURNAL J. Biol. Chem. 264 (21), 12573-12581 (1989)  
MEDLINE 89308691  
PUBMED 2745459  
COMMENT Original source text: Oryza sativa (strain M201, sub-species  
japonica) 2-week old germinating seed coleoptile tissue DNA.  
LOCATION/Qualifiers 1. .2774  
source  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="genomic DNA"  
/strain="M201"  
/sub\_species="japonica"  
/db\_xref="taxon:3947"  
/tissue\_type="coleoptile tissue"  
/aev\_stage="2-week old germinating seed"  
/gene="Gt1"  
709..718  
/gene="Gt1"  
741..1110  
/gene="Gt1"  
/number=1  
join(780..1110,1200..1474,1578..2057,2141..2554)  
/gene="Gt1"  
join(780..1110,1200..1474,1578..2057,2141..2554)  
/gene="Gt1"  
/codon\_start=1  
/product="glutelin"  
/protein\_id="AA50315.1"  
/db\_xref="GI:356399"  
/translation="WASINRPVFEFTVCLFLCDLSLAQQLLGSTSQMSRRSGSPR  
GCRFDLQAFERISVRSQAGTEFPDVNENLFOCTGVSVRRVIERGILLPHYNG  
ASVYLIIGRGITGTPGCEPEYQOQFOOSQALDPPRODELNGNRKPOAYRR  
GVIALPDAVAVKCYNDGEVVAIVTDINNGAOLDPPRODELNGNRKPOAYRR  
EVERSONTFSEFSELISRFGISNOVARLOLCONDOGRGIVVERGSLIDPVASL  
OREQOQMSRHHVQEGYCOOSYGSGCNGCDEPFRCMRKROITDNPNAALPYNRA  
GRVNLNSCFILMLVOMSVKVNLOHALSPWVNHASIVYITGSAOVQVYNN  
NGRTVNGELRGQLLVPOHYVVKRKAQREGCAVIAKRTNPNMVSIIAGKSLFRA  
LPPDYLANVLRISREARLKHNRGDEFGAFTPIQKSYQDVYVNAES"

intron  
/gene="Gt1"  
/number=1  
1200..2057  
/gene="Gt3"  
1200..1474  
/gene="Gt3"  
1475..1577  
/gene="Gt3"  
/number=2  
1578..2057  
/gene="Gt3"  
2058..2140  
/gene="Gt1"  
2141..2690  
/number=3  
/gene="Gt1"  
2582..2599  
/gene="Gt1"  
2690  
exon  
polyA\_signal  
/number=4  
2582..2599  
/gene="Gt1"  
2690  
polyA\_site  
/gene="Gt1"  
BASE COUNT 889 a 582 c 553 g 750 t  
ORIGIN



QY 61 AATATCATGTAATGTGTGAGCATTTAATGGACCAAGAAATAAAGAAACATTTTGTATG 120  
Db 217 AATATCATGTAATGTGTGAGCATTTAATGGACCAAGAAATAAAGAAACATTTTGTATG 276  
QY 121 GTGCTGATCTCGATGAGACCTCAAAAGTTCTCTCACCCGGATAGAAACCTTTAAGCA 180  
Db 277 GTGCTGATCTCGATGAGACCTCAAAAGTTCTCTCACCCGGATAGAAACCTTTAAGCA 336  
QY 181 ATGTGCAAAAGTTTGCATTTCTCCACTGACATTAATGCAAAATAAGATATATCAGATAGA 240  
Db 337 ATGTGCAAAAGTTTGCATTTCTCCACTGACATTAATGCAAAATAAGATATATCAGATAGA 396  
QY 241 GCAACTGATGATCATATATCATGAGCTCTCTCAACCTATATCTTCTTCTCATCTCATATAG 300  
Db 397 GCAACTGATGATCATATATCATGAGCTCTCTCAACCTATATCTTCTTCTCATCTCATATAG 456  
QY 301 TATCTTCAGCTAAATGTGTAGAAACATTAACCCATTAAGCAAGTTTATGATAGAGCGT 360  
Db 457 TATCTTCAGCTAAATGTGTAGAAACATTAACCCATTAAGCAAGTTTATGATAGAGCGT 516  
QY 361 GACACATGACAAATTCACAGACTCAAGCAAGATTAAGCAAAATGATGTGACATTAACACTC 420  
Db 517 GACACATGACAAATTCACAGACTCAAGCAAGATTAAGCAAAATGATGTGACATTAACACTC 576  
QY 421 CAGAGCTATATGTGATATGCAAAAGAGAGAGCTTATTAAGCAAGCAAGCAAGCAAGCAAG 480  
Db 577 CAGAGCTATATGTGATATGCAAAAGAGAGAGCTTATTAAGCAAGCAAGCAAGCAAGCAAG 636  
QY 481 AATTTCACTTGCCTTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 540  
Db 637 AATTTCACTTGCCTTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 696  
QY 541 AAAAGAAAG 600  
Db 697 AAAAGAAAG 756  
QY 601 CATATTCATCCACCTTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 660  
Db 757 CATATTCATCCACCTTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 816  
QY 661 GGACATTAACAACTGATATCTTAACATTAAGTCAAGAGCTTATCTCACTATTAATG 720  
Db 817 GGACATTAACAACTGATATCTTAACATTAAGTCAAGAGCTTATCTCACTATTAATG 876  
QY 721 CACAGTATTTTCATTTGTTTCTCAAAAAGCAATTAAGTCAATTAAGTCAATTAAGTCA 779  
Db 877 CACAGTATTTTCATTTGTTTCTCAAAAAGCAATTAAGTCAATTAAGTCAATTAAGTCA 935  
RESULT 4  
120872 120872 839 bp DNA linear PAT 07-OCT-1996  
LOCUS Sequence 1 from patent US 5516668.  
ACCESSION 120872  
VERSION 120872.1 GI:1601227  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
AUTHORS 1 (bases 1 to 839)  
TITLE Method for decreasing seed storage proteins and for transforming  
JOURNAL plants  
PATENT: US 5516668-A 1 14-MAY-1996;  
FEATURES Location/Qualifiers  
source 1..839  
BASE COUNT 307 a 170 c 135 g 227 t  
ORIGIN  
Query Match 98.0%; Score 763.4; DB 6; Length 839;  
Best Local Similarity 99.6%; Pred. No. 2,86-162;  
Matches 776; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

QY 1 TTCTGTATGACAGCAAAACCTTAATGAAGAGATGTGCTTTAGAAAAGCAAC 60  
Db 63 TTCTGTATGACAGCAAAACCTTAATGAAGAGATGTGCTTTAGAAAAGCAAC 122  
QY 61 AATTCATGTAATGTGTGAGCATTTAATGGACCAAGAAATAAAGAAACATTTTGTATG 120  
Db 123 AATTCATGTAATGTGTGAGCATTTAATGGACCAAGAAATAAAGAAACATTTTGTATG 182  
QY 121 GTGCTGATCTCGATGAGACCTCAAAAGTTCTCTCACCCGGATAGAAACCTTTAAGCA 180  
Db 183 GTGCTGATCTCGATGAGACCTCAAAAGTTCTCTCACCCGGATAGAAACCTTTAAGCA 242  
QY 181 ATGTGCAAAAGTTTGCATTTCTCCACTGACATTAATGCAAAATAAGATATATCAGATAGA 240  
Db 243 ATGTGCAAAAGTTTGCATTTCTCCACTGACATTAATGCAAAATAAGATATATCAGATAGA 302  
QY 241 GCAACTGATGATCATATATCATGAGCTCTCTCAACCTATATCTTCTTCTCATCTCATATAG 300  
Db 303 GCAACTGATGATCATATATCATGAGCTCTCTCAACCTATATCTTCTTCTCATCTCATATAG 362  
QY 301 TATCTTCAGCTAAATGTGTAGAAACATTAACCCATTAAGCAAGTTTATGATAGAGCGT 360  
Db 363 TATCTTCAGCTAAATGTGTAGAAACATTAACCCATTAAGCAAGTTTATGATAGAGCGT 422  
QY 361 GACACATGACAAATTCACAGACTCAAGCAAGATTAAGCAAAATGATGTGACATTAACACTC 420  
Db 423 GACACATGACAAATTCACAGACTCAAGCAAGATTAAGCAAAATGATGTGACATTAACACTC 482  
QY 421 CAGAGCTATATGTGATATGCAAAAGAGAGAGCTTATTAAGCAAGCAAGCAAGCAAGCAAG 480  
Db 483 CAGAGCTATATGTGATATGCAAAAGAGAGAGCTTATTAAGCAAGCAAGCAAGCAAGCAAG 542  
QY 481 AATTTCACTTGCCTTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 540  
Db 543 AATTTCACTTGCCTTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 602  
QY 541 AAAAGAAAG 600  
Db 603 AAAAGAAAG 662  
QY 601 CATATTCATCCACCTTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 660  
Db 663 CATATTCATCCACCTTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 720  
QY 661 GGACATTAACAACTGATATCTTAACATTAAGTCAAGAGCTTATCTCACTATTAATG 720  
Db 721 GGACATTAACAACTGATATCTTAACATTAAGTCAAGAGCTTATCTCACTATTAATG 780  
QY 721 CACAGTATTTTCATTTGTTTCTCAAAAAGCAATTAAGTCAATTAAGTCAATTAAGTCA 779  
Db 781 CACAGTATTTTCATTTGTTTCTCAAAAAGCAATTAAGTCAATTAAGTCAATTAAGTCA 839  
RESULT 5  
120873 120873 1240 bp DNA linear PAT 07-OCT-1996  
LOCUS Sequence 2 from patent US 5516668.  
ACCESSION 120873  
VERSION 120873.1 GI:1601228  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
AUTHORS 1 (bases 1 to 1240)  
TITLE Method for decreasing seed storage proteins and for transforming  
JOURNAL plants  
PATENT: US 5516668-A 2 14-MAY-1996;  
FEATURES Location/Qualifiers  
source 1..1240  
BASE COUNT 407 a 254 c 225 g 354 t





variation /note="t in lambda Ocg15; taa in [1]" 2283 /note="t in lambda Ocg15; c in [1]" 2340. 2345 join(2407. 2737, 2827. 3101, 3205. 3684, 3768. 4181) /codon\_start=1 /product="prepro-glute1in" /protein\_id="BA00462.1" /db\_xref="GI:218165" /translation="MASINRPIVETVTCFLICDGLAQQLIGSTSONMSSRSGSPR GGPRLDAPRIRSVSOGATTEFFDVSNELFOCTGTSYVRRIEPRRILLPHYTNK ASKVIIDGRIGTIPPGCPETVQOOFQSGAQLTSSQSKEDKHQIHFRR GDVIALPAGVAMWCYNDGEVVAIVYDINNGAQLDPRRQPLLAGNKRNPQYRR EYEMSONTFSGESTELLSEAFGISNQVARQLCCONDGRGIYVERGLSLQPYASL QEOEOQOMSRREHYEGGYOQSOFGSGCNGDETFCTMARVROMIDNPADYPRR GRVTLNTPONFPIILNYOMSAVKNLYONALLSPMTNASHIYITOGRAOVYNN NKTIVNGEIRRGILLIYPOHYVVKKAKRRECACTIAKTNPNMSTHAGSSIFRA LPIDVILANAKRISREBRQRLKHNKDEFGATFPDQTKSYQVIVAVSS" <2407. 2737 /number=1 /number=1 2738. 2826 /number=1 2827. 3101 /number=2 3102. 3204 /number=2 3128 /note="t in lambda Ocg15; c in [1]" 3205. 3684 /number=3 3489 /note="a in lambda Ocg15; t in lambda Ocg17" 3557 /note="a in lambda Ocg15; t in lambda Ocg [1,11]" 3623 /note="c in lambda Ocg15; g in [1]" 3685. 3767 /number=3 3768. >4181 /number=4 4042 /note="a in lambda Ocg15; c in lambda Ocg11" 4154 /note="c in lambda Ocg15; g in lambda Ocg17" 4209. 4226 /note="a in lambda Ocg15; g in lambda Ocg17" 4235. 4239 /note="poly(A) site consensus sequence common in higher plant genes" 4323. 4329 /note="poly(A) site consensus sequence common in higher plant genes" 4438. 4442 /note="poly(A) site consensus sequence common in higher plant genes" 4553 /note="g in lambda Ocg15; a in [1]" 4716 /note="g in lambda Ocg15; deletion in [1]" 4828 /note="a in lambda Ocg15; deletion in [1]" BASE COUNT 1652 a 1109 c 1074 g 1659 t ORIGIN Query Match 97.8%; Score 761.8; DB 8; Length 5494; Best local Similarity 99.5%; Pred. No. 5,3e-162; Matches 775; Conservative 0; Mismatches 2; Indels 2; Gaps 1; QY 1 TTCTGTAGTACAGACAACTAAATGATGAAAGAGATGGTGTAGAAAAGGAAC 60 Db 1630 TTCTGTAGTACAGACAACTAAATGATGAAAGAGATGGTGTAGAAAAGGAAC 1689 QY 61 AATATCATGAGTAAATGTGAGCATTAAGGACCAACGAATATAAAGAACTTTGTGTA 120

Db 1690 AATATCATGAGTAAATGTGAGCATTAAGGACCAACGAATATAAAGAACTTTGTGTA 1749 QY 121 GTCTGTAGTACAGACAACTAAATGATGAAAGAGATGGTGTAGAAAAGGAAC 180 Db 1750 GTCTGTAGTACAGACAACTAAATGATGAAAGAGATGGTGTAGAAAAGGAAC 1809 QY 181 ATGTGAAAGTGTGATTCCTCCACTGACATTAATGCAAAATAGATATGATGACATTA 240 Db 1810 ATGTGAAAGTGTGATTCCTCCACTGACATTAATGCAAAATAGATATGATGACATTA 1869 QY 241 GCAACTATGATCATATATGATGATGATGATGATGATGATGATGATGATGATGATG 300 Db 1870 GCAACTATGATCATATATGATGATGATGATGATGATGATGATGATGATGATGATG 1929 QY 301 TATCTAGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360 Db 1930 TATCTAGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1989 QY 361 GACACATGACAAATCAGACTCAAGACTCAAGACTCAAGACTCAAGACTCAAGACTC 420 Db 1990 GACACATGACAAATCAGACTCAAGACTCAAGACTCAAGACTCAAGACTCAAGACTC 2049 QY 421 CAGAGCTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480 Db 2050 CAGAGCTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2109 QY 481 AATTCATGCTGCTTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 540 Db 2110 AATTCATGCTGCTTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 2169 QY 541 AAAAGAAAG 600 Db 2170 AAAAGAAAG 2229 QY 601 CATATTCATCCACCTTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 660 Db 2230 CATATTCATCCACCTTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 2287 QY 661 GGACATTAACAACTATATGATGATGATGATGATGATGATGATGATGATGATGATG 720 Db 2288 GGACATTAACAACTATATGATGATGATGATGATGATGATGATGATGATGATGATG 2347 QY 721 CACGATATTTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 779 Db 2348 CACGATATTTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2406 RESULT 7 AC021891/c 144593 bp DNA linear PLN 09-AUG-2001 LOCUS Genomic Sequence for Oryza sativa, Nipponbare Strain, Chromosome X, Clone OSJNB40050N08, complete sequence. ACCESSION AC021891 VERSION AC021891.6 GI:12331454 KEYWORDS HTG. Oryza sativa Oryza sativa SOURCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza. REFERENCE 1 (bases 1 to 144593) Spiegel, L.A., Huang, E.N., Rodriguez, M.A., de la Bastide, M., Preston, R.R., Nascimento, L.U., King, L., Kirchoff, K.A., Vil, M.D., Baker, J.P., Miller, B., Toth, K., Shah, R.S., Bal, H., O'Shaughnessy, A., Dedhia, N.N. and McCombie, W.R. Genomic Sequence for Oryza sativa, Nipponbare Strain, Chromosome X, Clone OSJNB40050N08, Complete Sequence Unpublished 2 (bases 1 to 144593) McCombie, W.R. Direct Submission Submitted (22-JAN-2000) Lita Annenberg Hazen Genome Center, Cold Spring Harbor Laboratories, 1, Bungtown Road, Cold Spring Harbor,

REFERENCE 3 (bases 1 to 144593)  
AUTHORS McCombie,W.R.  
TITLE Direct Submission  
JOURNAL Submitted (22-JAN-2001) Lita Annenberg Hazen Genome Center, Cold Spring Harbor Laboratories, 1, Bungtown Road, Cold Spring Harbor, NY 11724, USA  
4 (bases 1 to 144593)  
REFERENCE Palmer,L.E., Bal,H., Spiegel,L.A., Huang,E.N., Rodriguez,M.A., de la Bastide,M., Preston,R.R., Nascimento,L.U., King,L., Kirchoff,K.A., Vil,M.D., Baker,J.P., Miller,B., Toch,K., Shah,R.S., O'Shaughnessy,A., Dedhia,N.N. and McCombie,W.R.  
AUTHORS  
TITLE Direct Submission  
JOURNAL Submitted (04-MAY-2001) Lita Annenberg Hazen Genome Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, NY 11724, USA  
REMARK Genomic Sequence For Oryza sativa, Nipponbare Strain, Chromosome X, Clone OSJNBa0050N08, complete sequence.  
REFERENCE 5 (bases 1 to 144593)  
AUTHORS Palmer,L.E., Bal,H., Spiegel,L.A., Huang,E.N., Rodriguez,M.A., de la Bastide,M., Preston,R.R., Nascimento,L.U., King,L., Kirchoff,K.A., Vil,M.D., Baker,J.P., Miller,B., Toch,K., Shah,R.S., O'Shaughnessy,A., Dedhia,N.N. and McCombie,W.R.  
TITLE Direct Submission  
JOURNAL Submitted (09-AUG-2001) Lita Annenberg Hazen Genome Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, NY 11724, USA  
REMARK Genomic Sequence For Oryza sativa, Nipponbare Strain, Chromosome X, Clone OSJNBa0050N08, complete sequence.  
COMMENT This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.  
FEATURES  
SOURCE  
Location/Qualifiers  
1..144593  
/organism="Oryza sativa"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:4530"  
/chromosome="10"  
/clone="OSJNBa0050N08"  
/clone\_1b="HindIII"  
/complement(234..4876)  
/gene="OSJNBa0050N08.1"  
/note="Putative retroelement"  
complement(join(234..696,1292..1522,1562..2076,2168..3695,3738..4876))  
/gene="OSJNBa0050N08.1"  
/note="Similar to putative polyprotein [Oryza sativa] (dbj|BAB39950.1) Expect = 0.0 identities = 928/1167 (79%)  
/codon\_start=1  
/product="Putative retroelement"  
/protein\_id="AAK50400.1"  
/db\_xref="GI:13940598"  
/translation="MAAQVPPVGVGPNAAQPHVAALSPFAIPYPOGAVNAGG  
EKGLGSGIKTRIPPOKPPVRYSGEDPEEFSTIESIAEAGDEMTAKYI  
HLALDGIARSMYFNLDNISYSEBOLRDVPLNRGYEPKIOOHLIGICESTI  
REVMRRSARCOVODITETASIVNAGLLEGRTPKANKEPIFTIHLICTIDGR  
RGEEDSKRQAIQAEYDKASTIAQAQOVAVAPPPLVSROSQPIQGQPPROSQAP  
ITWKFRTDRAGAAVAIVEVALKEEDAQAASHQQPHKRVKRDLYCAFHRSRH  
TTEQCRNIRORGNANRPQOGTVEAREBAVBOCPAPORODVHLSKRSKMQQR  
MWSHTSAGAGAPQIYNLOLISFPGDAGVTFPHODPLVISAIGFEYRRLVDGSI  
SADVFAEAYAKGGLPQALTFPAPYSLSGFGSEAVVLGGTILLIAFGSGNRREED  
LEFDVIDPYVNNALFGRAITLNKEFAISHNGLIKIKMPGPAAYVYVKGQSPATSGDL  
AIIIRHAYSVETPRHEKAKMPKTPPSKATTVVIDVPTKVISLGGMGSEJESI  
LEVIKKNIDIFAWSPDEVGVPANLIMHNLAVKPDIKPRQKLRKNSADROBAKAEV  
QKLRAEYIOEINHPEMLANPVLVRSNGKRWKCVDTDLNKAQCPKDFPPARIDQV  
DSTARCELMSEFLDAYSGHOIHMLNLDIPKTSFTTPGTGCHLIMPGILRNAGATPPT  
LVYKVLGKQGRNVEAYIDDIYVKSRAKFAHDALDQETPDSLAAGIKLNKEKCVFY

repeat\_region  
/note="Retrosat1 Ty3-Gypsy-retroelement from O.sativa"  
complement(1615..3980)  
repeat\_region  
/note="retro06 Oryza sativa gypsy-type retrotransposo"  
complement(3176..3904)  
/note="Class\_351"  
gene  
/complement(6732..7217)  
/gene="OSJNBa0050N08.2"  
/note="Putative disease resistance protein"  
complement(6732..7217)  
/gene="OSJNBa0050N08.2"  
/note="Similar to Arabidopsis thaliana protein similar to disease resistance response protein (dbj|BAB10492.1) Expect = 5e-11 identities = 34/81 (41%)  
/codon\_start=1  
/product="Putative disease resistance protein"  
/protein\_id="AAK50401.1"  
/db\_xref="GI:13940599"  
/translation="MAASHRAITALLILFSSSAATTIAAQAANADGGRTILHFMHE  
FENGPGNSMARVAPPARSGSLFEGVSVVDALREBADSATVRAQAAGASIAL  
DGSYTMLEPVTGDRRAARRRASGCRPAARSGMSAARARSQWREGRTMSAG  
R"  
7979..8029  
/note="Class\_1"  
complement(9162..9328)  
/note="RCB1a Oryza sativa DNA, centromere sequence R"  
complement(9333..9455)  
/note="RIRB3 Oryza sativa gypsy-type retrotransposon"  
8625..9923  
/note="Micropon4 Oryza sativa gene, repeat sequence M"  
complement(11458..13138)  
/gene="OSJNBa0050N08.3"  
/note="Putative retroelement"  
complement(join(11458..11712,11761..11853,11946..12118,12120..12478,12595..12864,13068..13138))  
/gene="OSJNBa0050N08.3"  
/note="Similar to putative Tam1 transposon protein TNP2 [Oryza sativa] (gb|AAK00419.1|AC069324.5) Expect = 1e-14 identities = 52/141 (36%)  
/codon\_start=1  
/product="Putative retroelement"  
/protein\_id="AAK50402.1"  
/db\_xref="GI:13940600"  
/translation="MLALIRALIEGYLFRRTROKRRRYNCLNRAPSLIGNDGPKLG  
SYRAQVIFRYLOPQASSTLIINFIYLPQPKVYPDAQSDGFSISQHRPLDPSSP  
NMMPSSVYNNOIDHEGCHTRIPRALHEGLRIGVGTKRKFYDDQYKVAHSSYL  
QGLATILRYTIDKRNKIRACNRPASATISGKQKRPFLKDONLSGSELEKXN  
AEAGRWPLPMDLMAVINGNYMFTFKRKDKSKSAQNSAYDGVDDWRATKTRAES  
KFLRLQPLNPRKTLVEXKTEIGEFVARRHDEDLALGTAHSGSKSTKRTTIN  
TLRKRAQTVQVQDFYRLAQAHPQAFPYLPQAEQVQVQIPSSIGSIETTSYDLIT  
GTPPEVTPYIGKA"  
complement(11246..12585)  
/note="transposon02 TNP2-like transposon protein gl14"  
complement(14322..15748)  
/note="transposon02 TNP2-like transposon protein gl14"  
complement(15221..16420)  
/gene="OSJNBa0050N08.4"  
/note="Putative retroelement"  
complement(join(15221..15484,15569..15994,16337..16420))  
/gene="OSJNBa0050N08.4"  
/note="Similar to putative retrotransposable elements TNP2 protein [Oryza sativa] (dbj|BAB16340.1) Expect = 4e-16 identities = 56/194 (28%)  
/codon\_start=1  
/product="Putative retroelement"



/protein\_id="AAP53625.1"  
/db\_xref="GI:31431919"  
/translation="MPVRSPLLSLPTGRRLRNGEARGEAGEARVPWRRAGQRRL  
PAPLSPPLDPADPARGGGGGGARGGAMTTHGVAAAPARLPSPLDLAGRG  
OGGGGARAGATTTRGAAAARLPSPPRGREGAGRRORRRGGEGGISAENVS  
MAGYAIRASSTNSAAGHPELRLSHRSPERTREIEQGLGRWVBDSG"  
complement(<4256..>6116)  
/locus\_tag="OSUNB0058B20.7"  
/note="Similar to (AP003208) hypothetical protein similar  
to Oryza sativa chromosome 1, P0416g11 19 E value 8e 15  
identities 52 267 (of subject)"  
complement(join(<4256..4458,4711..5342,6109..>6116))  
/locus\_tag="OSUNB0058B20.7"  
complement(join(11934..12425,12546..12736,12924..13222,  
13262..13807,13983..15376,15620..16576))  
/locus\_tag="OSUNB0058B20.5"  
complement(join(11934..12425,12546..12736,12924..13222,  
13262..13807,13983..15376,15620..16576))  
/locus\_tag="OSUNB0058B20.5"  
/codon\_start=1  
/product="putative retroelement"  
/protein\_id="AAP53628.1"

gene  
complement(<11934..>16576)  
/locus\_tag="OSUNB0058B20.5"  
/note="Similar to (AC093181) putative polyprotein (Oryza  
sativa) E value 0 0 identities 1089 1931 (of subject);  
contains pfam domain rve Integrase core domain e value 3  
e 33; contains pfam domain rvt Reverse transcriptase (RNA  
dependent e value 6 5e 29; contai"  
complement(join(<11934..12425,12546..12736,12924..13222,  
13262..13807,13983..15376,15620..>16576))  
/locus\_tag="OSUNB0058B20.5"  
complement(join(11934..12425,12546..12736,12924..13222,  
13262..13807,13983..15376,15620..16576))  
/locus\_tag="OSUNB0058B20.5"  
/codon\_start=1  
/product="putative retroelement"  
/protein\_id="AAP53628.1"

gene  
complement(<10984  
/locus\_tag="OSUNB0058B20.6"  
/note="hypothetical protein"  
join(<6199..6342,6458..6535,7355..7504,7737..8256,  
8513..8849,8903..9326,9491..9862,10013..10252,  
10346..10387,10547..>10984)  
/locus\_tag="OSUNB0058B20.6"  
join(6199..6342,6458..6535,7355..7504,7737..8256,  
8513..8849,8903..9326,9491..9862,10013..10252,  
10346..10387,10547..10984)  
/locus\_tag="OSUNB0058B20.6"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="AAP53627.1"  
/db\_xref="GI:31431921"  
/translation="MDISRTNEGYSCGQPVENSMWRAGVLLGASCLPYGPVAVG  
GVSMTYSDOSITYSNSRALYGAEOCTOHLTPQAPRPFSYFAAANAPITMAKRP  
NPASAKGPDGRVDNDNTAFILGFGSLSTGAGLPGLSLTPSSAPYRHRHEIAVIT  
PAGTKVPSVSNPNRPERSDIMPVNAASKDMMHTIRMPYHPIPEAGSDPAKI  
KCRRAIAPNRKPTIVDGNAMBARVILKRYCHSHEDVFEFCMLIFPLSSKQY  
ITVQGEVDLPMILPBGANPSSKRRKQCVKLAPRRKRAPASSDADDEDAER  
DGEERGEEREENVADKANAEVEDRIDPTGYPTPSPHNPTGVSSSPLRTDLE  
GAKALVAFSSGKAKEVRDSDGPTSPAGSLDLSLTAPIPLVPGAAGSIAAGA  
SASAEITVASAVFSGPLRELAIISPYKAKGCAGAKASASEYSLAHPHAPDEPT  
RADIVPEGVSNLVPAGTPELFTENPEDEGSAIKSLAIVRPSGGGRNKRQELG  
GGGGGOTLKAKAKAEVRLQSELDKCKEKAIVRYRIKAELEFKKASATVIVYTLTP  
KMEARLKASRABASAVEBSQFLEKMAKTESKTRCQTLRLALTDGTFEGCEHYA  
EFTYARKGDWEISAONIISPALNARIQFQWKGDSAAKRLLEOLAKAADRGEGD  
AACEEGGNADHDEAVGNPNPYCSRRRLGVHRSCIRPPSGSMRQSEKFAVKRGD  
KVRCPFRPVREMEINGRSYKVFVADMLRAHGRTLEDYDHYHRRLEDDARFMWH  
QKTRKETIFQRYSILVCSPPPLTKVYDIISSDESSPDHSLARDSDYGGEDVITL  
"

CDS  
complement(<11934..>16576)  
/locus\_tag="OSUNB0058B20.5"  
/note="Similar to (AC093181) putative polyprotein (Oryza  
sativa) E value 0 0 identities 1089 1931 (of subject);  
contains pfam domain rve Integrase core domain e value 3  
e 33; contains pfam domain rvt Reverse transcriptase (RNA  
dependent e value 6 5e 29; contai"  
complement(join(<11934..12425,12546..12736,12924..13222,  
13262..13807,13983..15376,15620..>16576))  
/locus\_tag="OSUNB0058B20.5"  
complement(join(11934..12425,12546..12736,12924..13222,  
13262..13807,13983..15376,15620..16576))  
/locus\_tag="OSUNB0058B20.5"  
/codon\_start=1  
/product="putative retroelement"  
/protein\_id="AAP53628.1"

Query Match

/db\_xref="GI:31431922"  
/translation="MAAQVPVGVGVGNNAQAOPHYOAIISPEATPPYPOQGAVNRAG  
EKCLPSGIGIKTRPIPPQFKPPVPYRSGTGPKEFLIYSALIAAHGCEANKAVI  
HLALDIAISWENLPDNSISYMSADIVNEGTVEERKTOOHLGIRQPGSI  
REYMRFSOARCOODITREASVIMASAGILGSELRLTKAKPELTLEHLICLIDFA  
RGSEBDRKROAIQATYDKASITTAQAQAQVYAEPPISVYOSQPAIQGPPQOSAP  
ITWKRTRIRAKGAVMAEEVQALRKREPEKQSNQOPARKKMKOTRWVHSITSG  
EGAPQYLNLIISFGEDEAEVIFPHODPLVISAELAGEVRRILIVDGGSSADVIFSA  
YAMKGTQALTPAPVSLRGFGGAQVQLOCTLLIARGSGNRREQLIFDVADIPY  
NYAINGRATLKPEAISHHNYLKLKMPGTAVIVYKGLQSPASKGDILAIRAHS  
VETPEHRKHKMPTPEKMTYQIDVDPKPTAISGDSGEGLESIIIVYKXID  
TEANSPDVGICPANTIMHILAVYDPDKPKCKLRMSADREAEKQVOKILRAVIT  
QELHPEMTLANVLEVRKSNQKRCVDFDNLNACXKDPFLPRIDIVDSTACELM  
SLFDAYSGYHQHNMPLDIPITSFTITPFGFDSIRAGIKLBNENAEARGLPFTKIG  
LGRVATYIDDIYVSKRAFDPAIDLEFPDLSIRAGIKLBNENAEARGLPFTKIG  
AGKFNTPPCQAAPELKOYLOSSPALISPPCSILLIYLAASPAVSAALVOETEG  
OKPYFTISELOGAKTRYIEMEKSLAIVWVSHKIKHYOAKHYVSAQVPLEIIRG  
KEYWGLKQWAEISPDRLHARAFKQVADPAEMTPALIMDEPDEOAGTAA  
ALSPGVPTIRYAAILOPTNNKTEFENILQIKRAKLAGYOLIRITDSKIYASHY  
DKSPKAEKEMRYLEAVRSMKCTGTIVYHLPDEAKRLQIRAKTKVSGQY  
RSGVLOPLKICISFAGEEMAKELHQGLCAHQAPVAKTIAVIMFPFWGIDIG  
PPIVARNKXKPAIVSEYFSGMIAEELGATSAAYOKVWNKIVICRGVYKEFIDN  
GKOPDSKPRKMEGELNLEIPPAVHAHQNGCAERTNCKILEALKRKBGAKKGW  
EELISYIMALWTPTRPSLARCFEMMRQ"  
complement(<18432..>18917)  
/locus\_tag="OSUNB0058B20.4"  
/note="Similar to (AC021891) putative disease resistance  
protein (Oryza sativa) E value 6e 48 identities 101 161  
(of subject); contains pfam domain disease\_resp Plant  
disease resistance response protel e value 9 1e 05"  
complement(<18432..>18917)  
/locus\_tag="OSUNB0058B20.4"  
complement(18432..18917)  
/locus\_tag="OSUNB0058B20.4"  
/codon\_start=1  
/product="putative disease resistance protein"  
/protein\_id="AAP53629.1"  
/db\_xref="GI:31431923"  
/translation="MAASHRALILILFSSSAATTAANOAPADOGSTTHLFYHME  
PFNGVPGSNARVAPRAGSGSLFQSVYDDLRGADPAIRLARGOGLAOTSILA  
DGSWTYMLDFPDGRTARAPWRSASGCHCPAARWRSGRMSANARASGMNEGTRSG  
R"  
complement(<20841..>21135)  
/locus\_tag="OSUNB0058B20.3"  
/note="unknown protein"  
complement(join(<20841..20929,20991..>21135))  
/locus\_tag="OSUNB0058B20.3"  
complement(join(20841..20929,20991..21135))  
/locus\_tag="OSUNB0058B20.3"  
/codon\_start=1  
/product="unknown protein"  
/protein\_id="AAP53630.1"  
/db\_xref="GI:31431924"  
/translation="MAAGPPPGGAAPPMPAAACHGRRPVGTGHPAGROMRGMAWK  
ERGGGRGEVEEDRGEDRKAKRGWRGIRPMW"  
complement(<23158..>24838)  
/locus\_tag="OSUNB0058B20.2"  
/note="Similar to (AC021891) putative retroelement (Oryza  
sativa) E value e 102 identities 207 406 (of subject))  
complement(join(<23158..23553,23818..24151,24768..>24838))  
/locus\_tag="OSUNB0058B20.2"  
complement(join(23158..23553,23818..24151,24768..24838))  
/locus\_tag="OSUNB0058B20.2"  
/codon\_start=1  
/product="putative retroelement"  
/protein\_id="AAP53631.1"  
/db\_xref="GI:31431925"  
/translation="MALIARAIEGYLFRRTROKREBRIPLPAHLEGRLIGKYGKKR  
FYDDYKAVAAASVYQOLATILEYTDENKKEKICACNDRSATWISKEOKHPPPL  
LKDNLISFGSLSEKNIARAGKMPMLPDLMASTIVKKEITGFPVARRSHDLITAALOT  
AHSGRVGRSSRTSMWVGFPESSSYKKRDTYNTLKRKKKTYQOVTOQFRIAAQHP  
"

gene  
complement(<20841..>21135)  
/locus\_tag="OSUNB0058B20.3"  
/note="unknown protein"  
complement(join(<20841..20929,20991..>21135))  
/locus\_tag="OSUNB0058B20.3"  
complement(join(20841..20929,20991..21135))  
/locus\_tag="OSUNB0058B20.3"  
/codon\_start=1  
/product="unknown protein"  
/protein\_id="AAP53630.1"  
/db\_xref="GI:31431924"  
/translation="MAAGPPPGGAAPPMPAAACHGRRPVGTGHPAGROMRGMAWK  
ERGGGRGEVEEDRGEDRKAKRGWRGIRPMW"  
complement(<23158..>24838)  
/locus\_tag="OSUNB0058B20.2"  
/note="Similar to (AC021891) putative retroelement (Oryza  
sativa) E value e 102 identities 207 406 (of subject))  
complement(join(<23158..23553,23818..24151,24768..>24838))  
/locus\_tag="OSUNB0058B20.2"  
complement(join(23158..23553,23818..24151,24768..24838))  
/locus\_tag="OSUNB0058B20.2"  
/codon\_start=1  
/product="putative retroelement"  
/protein\_id="AAP53631.1"  
/db\_xref="GI:31431925"  
/translation="MALIARAIEGYLFRRTROKREBRIPLPAHLEGRLIGKYGKKR  
FYDDYKAVAAASVYQOLATILEYTDENKKEKICACNDRSATWISKEOKHPPPL  
LKDNLISFGSLSEKNIARAGKMPMLPDLMASTIVKKEITGFPVARRSHDLITAALOT  
AHSGRVGRSSRTSMWVGFPESSSYKKRDTYNTLKRKKKTYQOVTOQFRIAAQHP  
"

gene  
complement(<20841..>21135)  
/locus\_tag="OSUNB0058B20.3"  
/note="unknown protein"  
complement(join(<20841..20929,20991..>21135))  
/locus\_tag="OSUNB0058B20.3"  
complement(join(20841..20929,20991..21135))  
/locus\_tag="OSUNB0058B20.3"  
/codon\_start=1  
/product="unknown protein"  
/protein\_id="AAP53630.1"  
/db\_xref="GI:31431924"  
/translation="MAAGPPPGGAAPPMPAAACHGRRPVGTGHPAGROMRGMAWK  
ERGGGRGEVEEDRGEDRKAKRGWRGIRPMW"  
complement(<23158..>24838)  
/locus\_tag="OSUNB0058B20.2"  
/note="Similar to (AC021891) putative retroelement (Oryza  
sativa) E value e 102 identities 207 406 (of subject))  
complement(join(<23158..23553,23818..24151,24768..>24838))  
/locus\_tag="OSUNB0058B20.2"  
complement(join(23158..23553,23818..24151,24768..24838))  
/locus\_tag="OSUNB0058B20.2"  
/codon\_start=1  
/product="putative retroelement"  
/protein\_id="AAP53631.1"  
/db\_xref="GI:31431925"  
/translation="MALIARAIEGYLFRRTROKREBRIPLPAHLEGRLIGKYGKKR  
FYDDYKAVAAASVYQOLATILEYTDENKKEKICACNDRSATWISKEOKHPPPL  
LKDNLISFGSLSEKNIARAGKMPMLPDLMASTIVKKEITGFPVARRSHDLITAALOT  
AHSGRVGRSSRTSMWVGFPESSSYKKRDTYNTLKRKKKTYQOVTOQFRIAAQHP  
"

gene  
complement(<20841..>21135)  
/locus\_tag="OSUNB0058B20.3"  
/note="unknown protein"  
complement(join(<20841..20929,20991..>21135))  
/locus\_tag="OSUNB0058B20.3"  
complement(join(20841..20929,20991..21135))  
/locus\_tag="OSUNB0058B20.3"  
/codon\_start=1  
/product="unknown protein"  
/protein\_id="AAP53630.1"  
/db\_xref="GI:31431924"  
/translation="MAAGPPPGGAAPPMPAAACHGRRPVGTGHPAGROMRGMAWK  
ERGGGRGEVEEDRGEDRKAKRGWRGIRPMW"  
complement(<23158..>24838)  
/locus\_tag="OSUNB0058B20.2"  
/note="Similar to (AC021891) putative retroelement (Oryza  
sativa) E value e 102 identities 207 406 (of subject))  
complement(join(<23158..23553,23818..24151,24768..>24838))  
/locus\_tag="OSUNB0058B20.2"  
complement(join(23158..23553,23818..24151,24768..24838))  
/locus\_tag="OSUNB0058B20.2"  
/codon\_start=1  
/product="putative retroelement"  
/protein\_id="AAP53631.1"  
/db\_xref="GI:31431925"  
/translation="MALIARAIEGYLFRRTROKREBRIPLPAHLEGRLIGKYGKKR  
FYDDYKAVAAASVYQOLATILEYTDENKKEKICACNDRSATWISKEOKHPPPL  
LKDNLISFGSLSEKNIARAGKMPMLPDLMASTIVKKEITGFPVARRSHDLITAALOT  
AHSGRVGRSSRTSMWVGFPESSSYKKRDTYNTLKRKKKTYQOVTOQFRIAAQHP  
"

gene  
complement(<20841..>21135)  
/locus\_tag="OSUNB0058B20.3"  
/note="unknown protein"  
complement(join(<20841..20929,20991..>21135))  
/locus\_tag="OSUNB0058B20.3"  
complement(join(20841..20929,20991..21135))  
/locus\_tag="OSUNB0058B20.3"  
/codon\_start=1  
/product="unknown protein"  
/protein\_id="AAP53630.1"  
/db\_xref="GI:31431924"  
/translation="MAAGPPPGGAAPPMPAAACHGRRPVGTGHPAGROMRGMAWK  
ERGGGRGEVEEDRGEDRKAKRGWRGIRPMW"  
complement(<23158..>24838)  
/locus\_tag="OSUNB0058B20.2"  
/note="Similar to (AC021891) putative retroelement (Oryza  
sativa) E value e 102 identities 207 406 (of subject))  
complement(join(<23158..23553,23818..24151,24768..>24838))  
/locus\_tag="OSUNB0058B20.2"  
complement(join(23158..23553,23818..24151,24768..24838))  
/locus\_tag="OSUNB0058B20.2"  
/codon\_start=1  
/product="putative retroelement"  
/protein\_id="AAP53631.1"  
/db\_xref="GI:31431925"  
/translation="MALIARAIEGYLFRRTROKREBRIPLPAHLEGRLIGKYGKKR  
FYDDYKAVAAASVYQOLATILEYTDENKKEKICACNDRSATWISKEOKHPPPL  
LKDNLISFGSLSEKNIARAGKMPMLPDLMASTIVKKEITGFPVARRSHDLITAALOT  
AHSGRVGRSSRTSMWVGFPESSSYKKRDTYNTLKRKKKTYQOVTOQFRIAAQHP  
"

gene  
complement(<20841..>21135)  
/locus\_tag="OSUNB0058B20.3"  
/note="unknown protein"  
complement(join(<20841..20929,20991..>21135))  
/locus\_tag="OSUNB0058B20.3"  
complement(join(20841..20929,20991..21135))  
/locus\_tag="OSUNB0058B20.3"  
/codon\_start=1  
/product="unknown protein"  
/protein\_id="AAP53630.1"  
/db\_xref="GI:31431924"  
/translation="MAAGPPPGGAAPPMPAAACHGRRPVGTGHPAGROMRGMAWK  
ERGGGRGEVEEDRGEDRKAKRGWRGIRPMW"  
complement(<23158..>24838)  
/locus\_tag="OSUNB0058B20.2"  
/note="Similar to (AC021891) putative retroelement (Oryza  
sativa) E value e 102 identities 207 406 (of subject))  
complement(join(<23158..23553,23818..24151,24768..>24838))  
/locus\_tag="OSUNB0058B20.2"  
complement(join(23158..23553,23818..24151,24768..24838))  
/locus\_tag="OSUNB0058B20.2"  
/codon\_start=1  
/product="putative retroelement"  
/protein\_id="AAP53631.1"  
/db\_xref="GI:31431925"  
/translation="MALIARAIEGYLFRRTROKREBRIPLPAHLEGRLIGKYGKKR  
FYDDYKAVAAASVYQOLATILEYTDENKKEKICACNDRSATWISKEOKHPPPL  
LKDNLISFGSLSEKNIARAGKMPMLPDLMASTIVKKEITGFPVARRSHDLITAALOT  
AHSGRVGRSSRTSMWVGFPESSSYKKRDTYNTLKRKKKTYQOVTOQFRIAAQHP  
"

gene  
complement(<20841..>21135)  
/locus\_tag="OSUNB0058B20.3"  
/note="unknown protein"  
complement(join(<20841..20929,20991..>21135))  
/locus\_tag="OSUNB0058B20.3"  
complement(join(20841..20929,20991..21135))  
/locus\_tag="OSUNB0058B20.3"  
/codon\_start=1  
/product="unknown protein"  
/protein\_id="AAP53630.1"  
/db\_xref="GI:31431924"  
/translation="MAAGPPPGGAAPPMPAAACHGRRPVGTGHPAGROMRGMAWK  
ERGGGRGEVEEDRGEDRKAKRGWRGIRPMW"  
complement(<23158..>24838)  
/locus\_tag="OSUNB0058B20.2"  
/note="Similar to (AC021891) putative retroelement (Oryza  
sativa) E value e 102 identities 207 406 (of subject))  
complement(join(<23158..23553,23818..24151,24768..>24838))  
/locus\_tag="OSUNB0058B20.2"  
complement(join(23158..23553,23818..24151,24768..24838))  
/locus\_tag="OSUNB0058B20.2"  
/codon\_start=1  
/product="putative retroelement"  
/protein\_id="AAP53631.1"  
/db\_xref="GI:31431925"  
/translation="MALIARAIEGYLFRRTROKREBRIPLPAHLEGRLIGKYGKKR  
FYDDYKAVAAASVYQOLATILEYTDENKKEKICACNDRSATWISKEOKHPPPL  
LKDNLISFGSLSEKNIARAGKMPMLPDLMASTIVKKEITGFPVARRSHDLITAALOT  
AHSGRVGRSSRTSMWVGFPESSSYKKRDTYNTLKRKKKTYQOVTOQFRIAAQHP  
"

gene  
complement(<20841..>21135)  
/locus\_tag="OSUNB0058B20.3"  
/note="unknown protein"  
complement(join(<20841..20929,20991..>21135))  
/locus\_tag="OSUNB0058B20.3"  
complement(join(20841..20929,20991..21135))  
/locus\_tag="OSUNB0058B20.3"  
/codon\_start=1  
/product="unknown protein"  
/protein\_id="AAP53630.1"  
/db\_xref="GI:31431924"  
/translation="MAAGPPPGGAAPPMPAAACHGRRPVGTGHPAGROMRGMAWK  
ERGGGRGEVEEDRGEDRKAKRGWRGIRPMW"  
complement(<23158..>24838)  
/locus\_tag="OSUNB0058B20.2"  
/note="Similar to (AC021891) putative retroelement (Oryza  
sativa) E value e 102 identities 207 406 (of subject))  
complement(join(<23158..23553,23818..24151,24768..>24838))  
/locus\_tag="OSUNB0058B20.2"  
complement(join(23158..23553,23818..24151,24768..24838))  
/locus\_tag="OSUNB0058B20.2"  
/codon\_start=1  
/product="putative retroelement"  
/protein\_id="AAP53631.1"  
/db\_xref="GI:31431925"  
/translation="MALIARAIEGYLFRRTROKREBRIPLPAHLEGRLIGKYGKKR  
FYDDYKAVAAASVYQOLATILEYTDENKKEKICACNDRSATWISKEOKHPPPL  
LKDNLISFGSLSEKNIARAGKMPMLPDLMASTIVKKEITGFPVARRSHDLITAALOT  
AHSGRVGRSSRTSMWVGFPESSSYKKRDTYNTLKRKKKTYQOVTOQFRIAAQHP  
"

gene  
complement(<20841..>21135)  
/locus\_tag="OSUNB0058B20.3"  
/note="unknown protein"  
complement(join(<20841..20929,20991..>21135))  
/locus\_tag="OSUNB0058B20.3"  
complement(join(20841..20929,20991..21135))  
/locus\_tag="OSUNB0058B20.3"  
/codon\_start=1  
/product="unknown protein"  
/protein\_id="AAP53630.1"  
/db\_xref="GI:31431924"  
/translation="MAAGPPPGGAAPPMPAAACHGRRPVGTGHPAGROMRGMAWK  
ERGGGRGEVEEDRGEDRKAKRGWRGIRPMW"  
complement(<23158..>24838)  
/locus\_tag="OSUNB0058B20.2"  
/note="Similar to (AC021891) putative retroelement (Oryza  
sativa) E value e 102 identities 207 406 (of subject))  
complement(join(<23158..23553,23818..24151,24768..>24838))  
/locus\_tag="OSUNB0058B20.2"  
complement(join(23158..23553,23818..24151,24768..24838))  
/locus\_tag="OSUNB0058B20.2"  
/codon\_start=1  
/product="putative retroelement"  
/protein\_id="AAP53631.1"  
/db\_xref="GI:31431925"  
/translation="MALIARAIEGYLFRRTROKREBRIPLPAHLEGRLIGKYGKKR  
FYDDYKAVAAASVYQOLATILEYTDENKKEKICACNDRSATWISKEOKHPPPL  
LKDNLISFGSLSEKNIARAGKMPMLPDLMASTIVKKEITGFPVARRSHDLITAALOT  
AHSGRVGRSSRTSMWVGFPESSSYKKRDTYNTLKRKKKTYQOVTOQFRIAAQHP  
"

gene  
complement(<20841..>21135)  
/locus\_tag="OSUNB0058B20.3"  
/note="unknown protein"  
complement(join(<20841..20929,20991..>21135))  
/locus\_tag="OSUNB0058B20.3"  
complement(join(20841..20929,20991..21135))  
/locus\_tag="OSUNB0058B20.3"  
/codon\_start=1  
/product="unknown protein"  
/protein\_id="AAP53630.1"  
/db\_xref="GI:31431924"  
/translation="MAAGPPPGGAAPPMPAAACHGRRPVGTGHPAGROMRGMAWK  
ERGGGRGEVEEDRGEDRKAKRGWRGIRPMW"  
complement(<23158..>24838)  
/locus\_tag="OSUNB0058B20.2"  
/note="Similar to (AC021891) putative retroelement (Oryza  
sativa) E value e 102 identities 207 406 (of subject))  
complement(join(<23158..23553,23818..24151,24768..>24838))  
/locus\_tag="OSUNB0058B20.2"  
complement(join(23158..23553,23818..24151,24768..24838))  
/locus\_tag="OSUNB0058B20.2"  
/codon\_start=1  
/product="putative retroelement"  
/protein\_id="AAP53631.1"  
/db\_xref="GI:31431925"  
/translation="MALIARAIEGYLFRRTROKREBRIPLPAHLEGRLIGKYGKKR  
FYDDYKAVAAASVYQOLATILEYTDENKKEKICACNDRSATWISKEOKHPPPL  
LKDNLISFGSLSEKNIARAGKMPMLPDLMASTIVKKEITGFPVARRSHDLITAALOT  
AHSGRVGRSSRTSMWVGFPESSSYKKRDTYNTLKRKKKTYQOVTOQFRIAAQHP  
"

gene  
complement(<20841..>21135)  
/locus\_tag="OSUNB0058B20.3"  
/note="unknown protein"  
complement(join(<20841..20929,20991..>21135))  
/locus\_tag="OSUNB0058B20.3"  
complement(join(20841..20929,20991..21135))  
/locus\_tag="OSUNB0058B20.3"  
/codon\_start=1  
/product="unknown protein"  
/protein\_id="AAP53630.1"  
/db\_xref="GI:31431924"  
/translation="MAAGPPPGGAAPPMPAAACHGRRPVGTGHPAGROMRGMAWK  
ERGGGRGEVEEDRGEDRKAKRGWRGIRPMW"  
complement(<23158..>24838)  
/locus\_tag="OSUNB0058B20.2"  
/note="Similar to (AC021891) putative retroelement (Oryza  
sativa) E value e 102 identities 207 406 (of subject))  
complement(join(<23158..23553,23818..24151,24768..>24838))  
/locus\_tag="OSUNB0058B20.2"  
complement(join(23158..23553,23818..24151,24768..24838))  
/locus\_tag="OSUNB0058B20.2"  
/codon\_start=1  
/product="putative retroelement"  
/protein\_id="AAP53631.1"  
/db\_xref="GI:31431925"  
/translation="MALIARAIEGYLFRRTROKREBRIPLPAHLEGRLIGKYGKKR  
FYDDYKAVAAASVYQOLATILEYTDENKKEKICACNDRSATWISKEOKHPPPL  
LKDNLISFGSLSEKNIARAGKMPMLPDLMASTIVKKEITGFPVARRSHDLITAALOT  
AHSGRVGRSSRTSMWVGFPESSSYKKRDTYNTLKRKKKTYQOVTOQFRIAAQHP  
"

gene  
complement(<20841..>21135)  
/locus\_tag="OSUNB0058B20.3"  
/note="unknown protein"  
complement(join(<20841..20929,20991..>21135))  
/locus\_tag="OSUNB0058B20.3"  
complement(join(20841..20929,20991..21135))  
/locus\_tag="OSUNB0058B20.3"  
/codon\_start=1  
/product="unknown protein"  
/protein\_id="AAP53630.1"  
/db\_xref="GI:31431924"  
/translation="MAAGPPPGGAAPPMPAAACHGRRPVGTGHPAGROMRGMAWK  
ERGGGRGEVEEDRGEDRKAKRGWRGIRPMW"  
complement(<23158..>24838)  
/locus\_tag="OSUNB0058B20.2"  
/note="Similar to (AC021891) putative retroelement (Oryza  
sativa) E value e 102 identities 207 406 (of subject))  
complement(join(<23158..23553,23818..24151,24768..>24838))  
/locus\_tag="OSUNB0058B20.2"  
complement(join(23158..23553,23818..24151,24768..24838))  
/locus\_tag="OSUNB0058B20.2"  
/codon\_start=1  
/product="putative retroelement"  
/protein\_id="AAP53631.1"  
/db\_xref="GI:31431925"  
/translation="MALIARAIEGYLFRRTROKREBRIPLPAHLEGRLIGKYGKKR  
FYDDYKAVAAASVYQOLATILEYTDENKKEKICACNDRSATWISKEOKHPPPL  
LKDNLISFGSLSEKNIARAGKMPMLPDLMASTIVKKEITGFPVARRSHDLITAALOT  
AHSGRVGRSSRTSMWVGFPESSSYKKRDTYNTLKRKKKTYQOVTOQFRIAAQHP  
"

gene  
complement(<20841..>21135)  
/locus\_tag="OSUNB0058B20.3"  
/note="unknown protein"  
complement(join(<20841..20929,20991..>21135))  
/locus\_tag="OSUNB0058B20.3"  
complement(join(20841..20929,20991..21135))  
/locus\_tag="OSUNB0058B20.3"  
/codon\_start=1  
/product="unknown protein"  
/protein\_id="AAP53630.1"  
/db\_xref="GI:31431924"  
/translation="MAAGPPPGGAAPPMPAAACHGRRPVGTGHPAGROMRGMAWK  
ERGGGRGEVEEDRGEDRKAKRGWRGIRPMW"  
complement(<23158..>24838)  
/locus\_tag="OSUNB0058B20.2"  
/note="Similar to (AC021891) putative retroelement (Oryza  
sativa) E value e 102 identities 207 406 (of subject))  
complement(join(<23158..23553,23818..24151,24768..>24838))  
/locus\_tag="OSUNB0058B20.2"  
complement(join(23158..23553,23818..24151,24768..24838))  
/locus\_tag="OSUNB0058B20.2"  
/codon\_start=1  
/product="putative retroelement"  
/protein\_id="AAP53631.1"  
/db\_xref="GI:31431925"  
/translation="MALIARAIEGYLFRRTROKREBRIPLPAHLEGRLIGKYGKKR  
FYDDYKAVAAASVYQOLATILEYTDENKKEKICACNDRSATWISKEOKHPPPL  
LKDNLISFGSLSEKNIARAGKMPMLPDLMASTIVKKEITGFPVARRSHDLITAALOT  
AHSGRVGRSSRTSMWVGFPESSSYKKRDTYNTLKRKKKTYQOVTOQFRIAAQHP  
"

gene  
complement(<20841..>21135)  
/locus\_tag="OSUNB0058B20.3"  
/note="unknown protein"  
complement(join(<20841..20929,20991..>21135))  
/locus\_tag="OSUNB0058B20.3"  
complement(join(20841..20929,20991..21135))  
/locus\_tag="OSUNB0058B20.3"  
/codon\_start=1  
/product="unknown protein"  
/protein\_id="AAP53630.1"  
/db\_xref="GI:31431924"  
/translation="MAAGPPPGGAAPPMPAAACHGRRPVGTGHPAGROMRGMAWK  
ERGGGRGEVEEDRGEDRKAKRGWRGIRPMW"  
complement(<23158..>24838)  
/locus\_tag="

Best Local Similarity 99.5%; Pred. No. 3.5e-152;		Matches 775; Conservative 0; Mismatches 2; Indels 2; Gaps 1;	
QY	1	TTCTGTAGTACAGACAAAATAAGTAAAGATGAGTGGTTTGAAGAAAGAAC	60
Db	95458	TTCTGTAGTACAGACAAAATAAGTAAAGATGAGTGGTTTGAAGAAAGAAC	95399
QY	61	AATATCATGAGTAATGTGTGAGCATTTTGGGACACGAAATTAAGAAACATTTTGATGA	120
Db	95398	AATATCATGAGTAATGTGTGAGCATTTTGGGACACGAAATTAAGAAACATTTTGATGA	95339
QY	121	GTCTGTATCTCTGAGTACGCTTCAAAAGTTCTTCCACCCGGGATTAAGAAACCTTTAAGCA	180
Db	95338	GTCTGTATCTCTGAGTACGCTTCAAAAGTTCTTCCACCCGGGATTAAGAAACCTTTAAGCA	95279
QY	181	ATGTGCAAAAGTTTGCATTCCTCCATGACATATATGCAAAATTAAGATTCATGAGACATA	240
Db	95278	ATGTGCAAAAGTTTGCATTCCTCCATGACATATATGCAAAATTAAGATTCATGAGACATA	95219
QY	241	GCAACTCATGACATATATGAGTGTCTCTCACTATTCATTCCTTACATCTACATGAAG	300
Db	95218	GCAACTCATGACATATATGAGTGTCTCTCACTATTCATTCCTTACATCTACATGAAG	95159
QY	301	TATCTTCAAGCTAAATGTGTGAGCATTTAAACCCATTAAGTACGTTGATGATTAAGGCT	360
Db	95158	TATCTTCAAGCTAAATGTGTGAGCATTTAAACCCATTAAGTACGTTGATGATTAAGGCT	95099
QY	361	GACACATGACAAATACACAGACTCAAGCAAGATTAAGCAAAATATGTGTACATAAAATC	420
Db	95098	GACACATGACAAATACACAGACTCAAGCAAGATTAAGCAAAATATGTGTACATAAAATC	95039
QY	421	CAGAGCTATATGTATATTTGCAAAAAGAGAGAGCTTATTAAGCAAGAGATACACACAA	480
Db	95038	CAGAGCTATATGTATATTTGCAAAAAGAGAGAGCTTATTAAGCAAGAGATACACACAA	94979
QY	481	AATTCACCTTCTCTTGTGTGCAAAAAGAGAGAGCTTATTAAGCAAGAGATACACACAA	540
Db	94978	AATTCACCTTCTCTTGTGTGCAAAAAGAGAGAGCTTATTAAGCAAGAGATACACACAA	94919
QY	541	AAAAGAAAAGAGAGAGAGACACACATGCTGCTCAATTAATATATCTGTATGATGCT	600
Db	94918	AAAAGAAAAGAGAGAGAGACACACATGCTGCTGCTCAATTAATATATCTGTATGATGCT	94859
QY	601	CATTATTCATCCACCTTCTGCTGACACACTTCATATATATTAAGAGTACCTGCTGCT	660
Db	94858	CATTATTCATCCACCTTCTGCTGACACACTTCATATATATTAAGAGTACCTGCTGCT	94801
QY	661	GGACATTAACAACCTATCTATTAAGTATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAG	720
Db	94800	GGACATTAACAACCTATCTATTAAGTATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAG	94741
QY	721	CACGATGATTTTCTCATTTGTTTCTCAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG	779
Db	94740	CACGATGATTTTCTCATTTGTTTCTCAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG	94682
RESULT 9			
LOCUS	RIGULTB	910 bp	DNA linear PLN 22-JAN-2003
DEFINITION	Oryza punctata gene for glutelin, 5' flanking region, clone:w08085.		
ACCESSION	D26364		
VERSION	D26364.1	GI:450478	
KEYWORDS	Oryza punctata		
SOURCE	Oryza punctata		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.		
REFERENCE	1. Takaiwa, F., Edinuma, H., Kikuchi, S. and Oono, K. Nucleotide sequence of a rice glutelin gene FEBS Lett. 221, 43-47 (1987).		
AUTHORS	TAKAIWA, F., EDINUMA, H., KIKUCHI, S. AND OONO, K.		
JOURNAL	FEBS Lett. 221, 43-47 (1987).		
REFERENCE	2		

Query Match		Score 729.4; DB 8; Length 910;	
Best Local Similarity 98.6%; Pred. No. 1.3e-154;		Matches 768; Conservative 0; Mismatches 6; Indels 5; Gaps 3;	
QY	1	TTCTGTAGTACAGACAAAATAAGTAAAGATGAGTGGTTTGAAGAAAGAAC	60
Db	137	TTCTGTAGTACAGACAAAATAAGTAAAGATGAGTGGTTTGAAGAAAGAAC	195
QY	61	AATATCATGAGTAATGTGTGAGCATTTTGGGACACGAAATTAAGAAACATTTTGATGA	120
Db	196	AATATCATGAGTAATGTGTGAGCATTTTGGGACACGAAATTAAGAAACATTTTGATGA	255
QY	121	GTCTGTATCTCTGAGTACGCTTCAAAAGTTCTTCCACCCGGGATTAAGAAACCTTTAAGCA	180
Db	256	GTCTGTATCTCTGAGTACGCTTCAAAAGTTCTTCCACCCGGGATTAAGAAACCTTTAAGCA	315
QY	181	ATGTGCAAAAGTTTGCATTCCTCCATGACATATATGCAAAATTAAGATTCATGAGACATA	240
Db	316	ATGTGCAAAAGTTTGCATTCCTCCATGACATATATGCAAAATTAAGATTCATGAGACATA	375
QY	241	GCAACTCATGACATATATGAGTGTCTCTCACTATTCATTCCTTACATCTACATGAAG	300
Db	376	GCAACTCATGACATATATGAGTGTCTCTCACTATTCATTCCTTACATCTACATGAAG	435
QY	301	TATCTTCAAGCTAAATGTGTGAGCATTTAAACCCATTAAGTACGTTGATGATTAAGGCT	360
Db	436	TATCTTCAAGCTAAATGTGTGAGCATTTAAACCCATTAAGTACGTTGATGATTAAGGCT	495
QY	361	GACACATGACAAATACACAGACTCAAGCAAGATTAAGCAAAATATGTGTACATAAAATC	420
Db	496	GACACATGACAAATACACAGACTCAAGCAAGATTAAGCAAAATATGTGTACATAAAATC	555
QY	421	CAGAGCTATATGTATATTTGCAAAAAGAGAGAGCTTATTAAGCAAGAGATACACACAA	480
Db	556	CAGAGCTATATGTATATTTGCAAAAAGAGAGAGCTTATTAAGCAAGAGATACACACAA	615
QY	481	AATTCACCTTCTCTTGTGTGCAAAAAGAGAGAGCTTATTAAGCAAGAGATACACACAA	540
Db	616	AATTCACCTTCTCTTGTGTGCAAAAAGAGAGAGCTTATTAAGCAAGAGATACACACAA	675
QY	541	AAAAGAAAAGAGAGAGAGACACACATGCTGCTCAATTAATATATCTGTATGATGCT	600
Db	676	AAAAGAAAAGAGAGAGAGACACACATGCTGCTGCTCAATTAATATATCTGTATGATGCT	733
QY	601	CATTATTCATCCACCTTCTGCTGACACACTTCATATATTAAGAGTACCTGCTGCT	660
Db	734	CATTATTCATCCACCTTCTGCTGACACACTTCATATATTAAGAGTACCTGCTGCT	791
QY	661	GGACATTAACAACCTATCTATTAAGTATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAG	720

AUTHORS Okita, T.W., Hwang, Y.S., Hnilo, J., Kim, W.T., Aryan, A.P., Larson, R. and Krishnan, H.B.  
 TITLE Structure and expression of the rice glutelin multigene family  
 JOURNAL J. Biol. Chem. 264 (21), 12573-12581 (1989)  
 MEDLINE 89308691  
 PUBMED 2745459  
 REFERENCE 3. (bases 1 to 910)  
 AUTHORS Wu, H.-K.  
 TITLE Direct Submission  
 JOURNAL Submitted (27-DEC-1993) Hsin-Kan Wu, Institute of Botany, Academia Sinica, Nankang, Taipei, Taiwan 11529, Republic of China  
 (E-mail: BOHSKAW@CCVAX.SINICA.EDU.TW, Tel:886-2-7899590(ex.112))

FEATURES  
 source  
 1. 910  
 /organism="Oryza punctata"  
 /mol\_type="genomic DNA"  
 /strain="W1564"  
 /db\_xref="taxon:4537"  
 /clone="w08085"  
 /tissue\_type="leaf"  
 /germline  
 BASE COUNT 328 a 174 c 155 g 253 t  
 ORIGIN

```
|||||
Db 792 GGAATTAAACAACTCTACCTTAACATTTAGATTCAGAGACCTTTATCTCTCCTTAATAG 851
QY 721 CAGCATGATTTCTCATTTGTTTCTCCAAAAAGCATTCAGTTCATTTAGTCCCTACACAAAC 779
852 CAGCATGATTTCTCATTTGTTTCTCCAAAAAGCATTCAGTTCATTTAGTCCCTACACAAAC 910

RESULT 10
RIGL1P 912 bp DNA linear PLN 22-JAN-2003
LOCUS Oryza grandiglumis gene for glutelin, 5' flanking region,
DEFINITION clone:w13095.
ACCESSION D26368.1 GI:450482
VERSION D26368.1
KEYWORDS Oryza grandiglumis
SOURCE Oryza grandiglumis
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE 1
AUTHORS Takaiwa,F., Ebinuma,H., Kikuchi,S. and Oono,K.
TITLE Nucleotide sequence of a rice glutelin gene
JOURNAL FEBS Lett. 221, 43-47 (1987)
REFERENCE 2
AUTHORS Okita,T.W., Hwang,Y.S., Hnilo,J., Kim,W.T., Aryan,A.P., Larson,R.
and Krishnan,H.B.
TITLE Structure and expression of the rice glutelin multigene family
JOURNAL J. Biol. Chem. 264 (21), 12573-12581 (1989)
MEDLINE 89308691
PUBMED 2745459
REFERENCE 3 (bases 1 to 912)
AUTHORS Wu,H.-K.
TITLE Direct Submission
JOURNAL Submitted (27-DEC-1993) Hsin-Kan Wu, Institute of Botany, Academia
Sinica, Nankang, Taipei, Taiwan 11529, Republic of China
(E-mail:BOHSKAW@CCVAX.SINICA.EDU.TW, Tel:886-2-7899590(ex.112))
Location/Qualifiers
FEATURES
source
1..912
/organism="Oryza grandiglumis"
/mol_type="genomic DNA"
/strain="w1194"
/db_xref="taxon:26960"
/clone="w13095"
/tissue_type="leaf"
/germline

BASE COUNT 329 a 175 c 156 g 252 t
ORIGIN
Query Match 93.6%; Score 729.4; DB 8; Length 912;
Best Local Similarity 98.6%; Pred. No.1.3e-154;
Matches 768; Conservative 0; Mismatches 6; Indels 5; Gaps 3;

QY 1 TTCTGTACTAAGACAAACTATAAGTATGAAAGAAAGTGTGTGTAAGAAAGGAAC 60
Db 139 TTCTGTACTAAGACAAACTATAAGTATGAAAGAAAGTGTGTGTAAGAAAGGAAC 197
QY 61 AATATCATGAGTATGCTGAGCATTTGAGACACAGAAATVAAAAGAAACATTTGANGA 120
Db 198 AATATCATGAGTATGCTGAGCATTTGAGACACAGAAATVAAAAGAAACATTTGANGA 257
QY 121 GTGTGTATCTCGATGAGCCTGAAAAGTCTCTGACCCCGGATAGAAAACCTTAAACA 180
Db 258 GTGTGTATCTCGATGAGCCTGAAAAGTCTCTGACCCCGGATAGAAAACCTTAAACA 317
QY 181 AGTGTGAAGTTGCATTTCCAGCTGACATTAATGAAATATGATTCATGATGACATA 240
Db 318 AGTGTGAAGTTGCATTTCCAGCTGACATTAATGAAATATGATTCATGATGACATA 377
QY 241 GGAAGTCATGATATATGATGCTCTGCAACCTATTCATTCCTACTCATATAG 300
Db 378 GGAAGTCATGATATATGATGCTCTGCAACCTATTCATTCCTACTCATATAG 437

|||||
QY 301 TATCTTACGCTAAATGTGTGAACATTAACCCATTAAGTCACGTTTGTAGATTTAGGCGT 360
Db 438 TATCTTACGCTAAATGTGTGAACATTAACCCATTAAGTCACGTTTGTAGATTTAGGCGT 497
QY 361 GAGACATGACAAATCCAGAGCTCAAGCAAGATTAAGCAAAATGATGTGACTTAAACTC 420
Db 498 GAGACATGACAAATCCAGAGCTCAAGCAAGATTAAGCAAAATGATGTGACTTAAACTC 557
QY 421 CAGAGCTATATGCTATATGTCACAAAAAGAGAGAGACTTATTAAGCAAGGACTGTCACAA 480
Db 558 CAGAGCTATATGCTATATGTCACAAAAAGAGAGAGACTTATTAAGCAAGGACTGTCACAA 617
QY 481 AATTCACCTTGCCTTTCGTGTGAAAAAGAGAGGCTTTACATTACATGTCATATATGC 540
Db 618 AATTCACCTTGCCTTTCGTGTGAAAAAGAGAGGCTTTACATTACATGTCATATATGC 677
QY 541 AAAAGAAAGAGAAAGAAACACACAGATGCTGCGCTCATATATCATATGCTATGCTCAT 600
Db 678 AAAAGAA--AGAGAAAGACACACAGATGCTGCGCTCATATATCATATGCTATGCTCAT 735
QY 601 CATATTCATCCACCTTTCGTGTACACACTTCATATATGATTAAGAGTCACTTCACGCT 660
Db 736 CATATTCATCCACCTTTCGTGTACACACTTCATATATGCT--GAGTCACTTCATGCT 793
QY 661 GGAATTTAAACAATCTATATCTTAACATTTAGTGAAGAGAGGCTTTATCTCATATTAATG 720
Db 794 GGAATTTAAACAATCTATATCTTAACATTTAGTGAAGAGAGGCTTTATCTCATATTAATG 853
QY 721 CAGCATGATTTCTCATTTGTTTCTCCAAAAAGCATTCAGTTCATTTAGTCCCTACACAAAC 779
Db 854 CAGCATGATTTCTCATTTGTTTCTCCAAAAAGCATTCAGTTCATTTAGTCCCTACACAAAC 912

RESULT 11
RIGL1P 910 bp DNA linear PLN 22-JAN-2003
LOCUS Oryza australiensis gene for glutelin, 5' flanking region,
DEFINITION clone:w14095.
ACCESSION D26369.1 GI:450483
VERSION D26369.1
KEYWORDS Oryza australiensis
SOURCE Oryza australiensis
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE 1
AUTHORS Takaiwa,F., Ebinuma,H., Kikuchi,S. and Oono,K.
TITLE Nucleotide sequence of a rice glutelin gene
JOURNAL FEBS Lett. 221, 43-47 (1987)
REFERENCE 2
AUTHORS Okita,T.W., Hwang,Y.S., Hnilo,J., Kim,W.T., Aryan,A.P., Larson,R.
and Krishnan,H.B.
TITLE Structure and expression of the rice glutelin multigene family
JOURNAL J. Biol. Chem. 264 (21), 12573-12581 (1989)
MEDLINE 89308691
PUBMED 2745459
REFERENCE 3 (bases 1 to 910)
AUTHORS Wu,H.-K.
TITLE Direct Submission
JOURNAL Submitted (27-DEC-1993) Hsin-Kan Wu, Institute of Botany, Academia
Sinica, Nankang, Taipei, Taiwan 11529, Republic of China
(E-mail:BOHSKAW@CCVAX.SINICA.EDU.TW, Tel:886-2-7899590(ex.112))
Location/Qualifiers
FEATURES
source
1..910
/organism="Oryza australiensis"
/mol_type="genomic DNA"
/strain="w0008"
/db_xref="taxon:4532"
/clone="w14095"
/tissue_type="leaf"
/germline
```

BASE COUNT 328 a 175 c 154 g 253 t  
 ORIGIN  
 Query Match 92.1%; Score 717.4; DB 8; Length 910;  
 Best Local Similarity 98.5%; Pred. No. 6.8e-152;  
 Matches 767; Conservative 0; Mismatches 6; Indels 6; Gaps 4;

QY 1 TTCTGTAGTACAGCAAACTAAAGTATATAAGAAAGATGTGTGTAGAAAAGAAAC 60  
 |||||  
 DB 138 TTCTGTAGTACAGCAAACT-AAAATATATAAGAAAGATGTGTGTAGAAAAGAAAC 196  
 QY 61 AATATCATGAGTATATGAGCATTTATGAGCAGCAGCAAGAAATATAAGCACTTTGATGA 120  
 |||||  
 DB 197 AATATCATGAGTATATGAGCATTTAT-GGACACAGCAAAATATAAGCACTTTGATGA 255  
 QY 121 GTGTGTATCTCTGAGAGAGCCCAAAAGTTCTCTCACCCGGATAGAAAACCTTAAGCA 180  
 |||||  
 DB 256 GTGTGTATCTCTGAGAGAGCCCAAAAGTTCTCTCACCCGGATAGAAAACCTTAAGCA 315  
 QY 181 ATGTGCAAAAGTTTGCATTTCCACTGACATTAATGCAAAATATAAGATATCATGATGACATA 240  
 |||||  
 DB 316 ATGTGCAAAAGTTTGCATTTCTCCACTGACATTAATGCAAAATATAAGATATCATGATGACATA 375  
 QY 241 GCAACTCATGACATTAATGATGCTCTCTCCACTTATTTCTCTCATCTCATCTCATTAAG 300  
 |||||  
 DB 376 GCAACTCATGACATTAATGATGCTCTCTCTCTCATCTTATTTCTCTCATCTCATTAAG 435  
 QY 301 TATCTTCAGCTAAATGTTAGAAATTAACCCATTAAGTACAGTTTGTAGATTAAGGCGT 360  
 |||||  
 DB 436 TATCTTCAGCTAAATGTTAGAAATTAACCCATTAAGTACAGTTTGTAGATTAAGGCGT 495  
 QY 361 GACACATGACAAATTCACAGACACAGCAAGATTAAGCAAAATATGTGTACATTAATCTC 420  
 |||||  
 DB 496 GACACATGACAAATTCACAGACACAGCAAGATTAAGCAAAATATGTGTACATTAATCTC 555  
 QY 421 CAGAGCTATATGTCATTTGCAAAAGAGAGGAGCTTATTAAGACAAAGCATACACACAA 480  
 |||||  
 DB 556 CAGAGCTATATGTCATTTGCAAAAGAGAGGAGCTTATTAAGACAAAGCATACACACAA 615  
 QY 481 AATTCACCTTGCCTTGTGTGTCACAAAGAGAGGAGCTTATTAAGTATCATTAATCATTTGC 540  
 |||||  
 DB 616 AATTCACCTTGCCTTGTGTGTCACAAAGAGAGGAGCTTATTAAGTATCATTAATCATTTGC 675  
 QY 541 AAAAGAAAGAGAAAGAAAGAAAGAAAGAGAGGAGCTTATTAAGTATCATTAATCATTTGC 600  
 |||||  
 DB 676 AAAAGAA--AGAGAAAGAAAGAAAGAGAGGAGCTTATTAAGTATCATTAATCATTTGC 733  
 QY 601 CATTTATCATTCACCTTGTGTGTCACACACTTCAATATATATATAGAGTCACTTCACGTCT 660  
 |||||  
 DB 734 CATTTATCATTCACCTTGTGTGTCACACACTTCAATATATCT--GAGTCACTTCATTTCT 791  
 QY 661 GGAACATTAACAACTCTATCTTAACATTTAGATGCAAGAGCCTTTATCTCACTATTAATG 720  
 |||||  
 DB 792 GGACATTAACAACTCTATCTTAACATTTAGATGCAAGAGCCTTTATCTCACTATTAATG 851  
 QY 721 CAGAGTATTCATTCATTTGTGTGTCACAAAGAGAGCTTATTAAGTATCATTTGCACAAAC 779  
 |||||  
 DB 852 CAGAGTATTCATTCATTTGTGTGTCACAAAGAGAGCTTATTAAGTATCATTTGCACAAAC 910

RESULT 12  
 RIGGT2A  
 LOCUS RIGGT2A 2867 bp DNA linear PLN 08-OCT-1994  
 DEFINITION Oryza sativa glutelin (Gt2) gene, complete cds.  
 ACCESSION L36819 M28157  
 VERSION L36819.1 GI:556402  
 KEYWORDS glutelin.  
 SOURCE Oryza sativa (japonica cultivar-group)  
 ORGANISM Oryza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.  
 REFERENCE 1 (sites)

AUTHORS Ohta, T.W., Hwang, Y.S., Hnilo, J., Kim, W.T., Aryan, A.P., Larson, R.  
 and Krishnan, H.B.  
 TITLE Structure and expression of the rice glutelin multigene family  
 JOURNAL J. Biol. Chem. 264 (21), 12573-12581 (1989)  
 MEDLINE 89308691  
 PUBMED 2745459  
 COMMENT Original source text: Oryza sativa (strain M201, sub-species japonica) 2-week old germinating seed coleoptile tissue DNA.  
 FEATURES  
 source  
 1..2867  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /mol\_type="genomic DNA"  
 /strain="M201"  
 /sub\_species="japonica"  
 /db\_xref="taxon:39947"  
 /tissue\_type="coleoptile tissue"  
 /dev\_stage="2-week old germinating seed"  
 765..775  
 /gene="Gt2"  
 815..820  
 /gene="Gt2"  
 843..1212  
 /gene="Gt2"  
 /number=1  
 join(882..1212,1302..1576,1682..2161,2245..2658)  
 /gene="Gt2"  
 join(882..1212,1302..1576,1682..2161,2245..2658)  
 /gene="Gt2"  
 /codon\_start=1  
 /product="glutelin"  
 /protein\_id="AAA50317.1"  
 /db\_xref="GI:556403"  
 /translation="MASTIRPIVEFTVQLFLNGSLAQLIGSTSGQSSRGSGR  
 ECRDRLOAEPIRSVSOAGTEFPVSNEDQCGVGVVRVLEPGILLPIRYTG  
 ASLVIYILOGRGITGPFPGCEPTVOQFOGSGAQLTESQSOQKFEDEHKIRFRQ  
 GDVALPAGAHMCYMDGEVPAIYVTDLNNAGNOLDPRODFLLAGNKRNPAYR  
 EVERSONIRSGPSTELSEALGVSGVARQLCCORDRGELIVRYEHLSTLOEYPSL  
 OEQOGQVGRERYPOGQSQSQSGCSGSGIDETCRKRVQNTDNRDPTNPPL  
 GRTVNLNTQNFPLNLVMSAKVNLTONLILSPFNINAHSAVITITGRRARVYNN  
 NGKIVFNGEPRRGQLIVQHVYVKKRQREGCAVLAATNNSVSHIAGKSSIFRA  
 LPNDVLNAYRIREDAQLKHNKRDGAFPIQYKSYQDYNAESS"  
 1213..1301  
 /gene="Gt2"  
 /number=1  
 1302..1576  
 /gene="Gt2"  
 /number=2  
 1577..1681  
 /gene="Gt2"  
 /number=2  
 1682..2161  
 /gene="Gt2"  
 /number=4  
 2162..2244  
 /gene="Gt2"  
 /number=3  
 2245..2768  
 /gene="Gt2"  
 /number=4  
 2685..2690  
 /gene="Gt2"  
 /note="putative"  
 2689..2694  
 /gene="Gt2"  
 /note="putative"  
 2768  
 /gene="Gt2"  
 polyA\_signal  
 polyA\_signal  
 polyA\_signal  
 polyA\_site  
 2768  
 /gene="Gt2"  
 BASE COUNT 942 a 593 c 564 g 768 t  
 ORIGIN  
 Query Match 50.3%; Score 392.2; DB 8; Length 2867;  
 Best Local Similarity 76.9%; Pred. No. 1.8e-78;  
 Matches 690; Conservative 0; Mismatches 73; Indels 134; Gaps 12;





LOCUS	RIGLND	481 bp	DNA	linear	PLN 22-JUN-2003
DEFINITION	Oryza eichingeri gene for glutelin, 5' flanking region,				
ACCESSION	D63966				
VERSION	D63966.1	GI:450480			
KEYWORDS					
SOURCE	Oryza eichingeri				
ORGANISM	Oryza eichingeri				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; Liliopsida; Poales; Poaceae; Ehrhartioideae; Oryzeae; Oryza.				
TITLE	1 Takaiwa, F., Ebinuma, H., Kikuchi, S. and Oono, K.				
JOURNAL	Nucleotide sequence of a rice glutelin gene				
AUTHORS	2 Press Lett. 221, 43-47 (1987)				
TITLE	Okita, T.W., Hwang, Y.S., Hillo, J., Kim, W.T., Aryan, A.P., Larson, R.				
JOURNAL	and Krishnan, H.B.				
MEDLINE	Structure and expression of the rice glutelin multigene family				
PUBMED	J. Biol. Chem. 264 (21), 12573-12581 (1989)				
REFERENCE	89308691				
AUTHORS	2745459				
TITLE	3 (bases 1 to 481)				
JOURNAL	Wu, H.-K.				
FEATURES	Direct Submission				
source	Submitted (27-DEC-1993) Hsin-Kan Wu, Institute of Botany, Academia Sinica; Nankang, Taipei, Taiwan 11529, Republic of China				
	(E-mail:BOHSHAW@CCVAX.SINICA.EDU.TW, Tel.:886-2-7899590(ex.112))				
	Location/Qualifiers				
	1..481				
	/organism="Oryza eichingeri"				
	/mol_type="genomic DNA"				
	/strain="W1519"				
	/db_xref="taxon:29689"				
	/clone="W11050"				
	/tissue_type="leaf"				
	/germline				
BASE COUNT	165 a 94 c 83 g 139 t				
ORIGIN					
Query Match	37.2%; Score 289.4; DB 8; Length 481;				
Best Local Similarity	90.68; Pred. NO. 3.se-55;				
Matches 356; Conservative 0; Mismatched 26; Indels 11; Gaps 4;					
OY	387 CAAGATTAAGCAAAATGATGTGTACATAAACAACCGACGCTATGTCATTGGCAAAA	446			
DB	100 CACCATTAAAGCAAAGAGTGTACAAAAAACGCCAGACTATGTCATTGGCAAAA	159			
OY	447 GAGGAGAGCTTTAABAACAGGCATGACTCACAAAATTCACTTGCTTGCTGAATA	506			
DB	160 GAGGAGAGCCTTAACACAGGAGCATGACTGAGAATAATATTGCTTGCTGTAACAAA	219			
OY	507 AGAGAGAGGCTTTACATTTCCATGTGATATTTGCAAAAGAAAGAGAGAAAGACAACA	566			
DB	220 AGAGAGAGC-----TTTTCATGTGCTATTTGCAAAA--AGAGAGAGACATTA	271			
OY	567 ATGCTGGTCAATTATACATATCTGTATGTCCATCTATTATTCATCCACTTTCGTATAC	626			
DB	272 ATGCTGGTCAATTATACATATCTGTAT-TCCATCATTTATTCATCCACTTTCGTATAC	330			
OY	627 ACATCTCATTTATCATTAAGACATCCAGTCCGAGCAATTAACAACCTCATCTTAACA	686			
DB	331 KCACCTCATTAATCAT--GAGTCATCTATGTGTGACATTAACAACCTCATCTTAACA	388			
OY	687 TTTAGATGCAAGAGCCTTATTCACATATAAAGCAAGATGATTTCTCATGTTCTGCAC	746			
DB	389 TTTAGATGCAAGAGCCTTATTCACATATAAAGCAAGATGATTTCTCATGTTCTGCAC	448			
OY	747 AAAAGCATTCAGTTCAATAGTCTCAACAAC 779				
DB	449 AAAAGCATTCAGTTCAATAGTCTCAACAAC 481				

RESULT 15	LOCUS	LOCUS	DNA	linear	PLN 22-JAN-2003
LOCUS	RIGGLTA	1119 bp			
DEFINITION	Oryza perennis gene for glutelin, 5' flanking region, clone:w05120.				
ACCESSION	D26363				
VERSION	D26363.1	GI:517152			
KEYWORDS					
SOURCE	Oryza perennis				
ORGANISM	Oryza perennis				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.				
REFERENCE	1				
AUTHORS	Takaiwa,F., Ebinuma,H., Kikuchi,S. and Oono,K.				
TITLE	Nucleotide sequence of a rice glutelin gene				
JOURNAL	FEBS Lett. 221, 43-47 (1987)				
REFERENCE	2				
AUTHORS	Oikita,T.W., Hwang,Y.S., Hnilo,J., Kim,W.T., Aryan,A.P., Larson,R. and Krishnan,H.B.				
TITLE	Structure and expression of the rice glutelin multigene family				
JOURNAL	J. Biol. Chem. 264 (21), 12573-12581 (1989)				
MEDLINE	89308681				
REFERENCE	274539				
AUTHORS	3 (bases 1 to 1119)				
TITLE	Wu,H.-K.				
JOURNAL	Direct Submission				
COMMENT	Submitted (27-DEC-1993) Hsin-Kan Wu, Institute of Botany, Academia Sinica, Nankang, Taipei, Taiwan 11529, Republic of China (E-mail:BOHSKANU@CCVAX.SINICA.EDU.TW, Tel:886-2-789590(ex.112))				
FEATURES	On Jul 30, 1994 this sequence version replaced gi:450477.				
FEATURES	Location/Qualifiers				
source	1. 1119				
	/organism="Oryza perennis"				
	/mol_type="genomic DNA"				
	/strain="W0107"				
	/db_xref="taxon:29691"				
	/clone="w05120"				
	/tissue_type="leaf"				
	/germline				
BASE COUNT	427 a 209 c 191 g 292 t				
ORIGIN					
Query Match	37.1%; Score 289.2; DB 8; Length 1119;				
Best Local Similarity	82.8%; Pred. No.3.6e-55;				
Matches 447; Conservative 0; Mismatches 63; Indels 30; Gaps 9;					
OY	1 TTCGTGTGTCAACAAACCTAAAGTAAGAAGAAGTGGGTAGTAAAGGAAC 60				
DB	140 TTCGTGTGTCAACAAACCTAAAGTAAGAAGAAGTGGGTAGTAAAGGAAC 196				
OY	61 AATATCATGAGTATGTGTGACCATTAATGGGACACAGAAAT-AAAAAGACATTTGAT 118				
DB	197 AATATTAATGAGTATGTGTGACCATTAATGGGACACAGAAATAAAAAGACATTTTAT 256				
OY	119 GAGTCGGTATTCGTCATGAGCCGTCACAAATTCCTCCACCGGGATAGAAACCTTAAG 178				
DB	257 GAGCAGTGTGTCTCATATAGCCCTTGATGTTCACACCGAGATAGAAACCTTTAG 316				
OY	179 C-----AATGTCAAAGTTTCATCTTCACATGACATATATGCAAA 218				
DB	317 CAATGAACAATGGAACCGTTTAATGTGGAAGTTGGCATTTCCGC- GACATATATGCAA 375				
OY	219 AATAGATATCATGATGACATATGCACTCATCATATATATAGCCCTCTCAACTAT 278				
DB	376 AAAAAATATATATTAAGATATGCAATGCAAGTCATGATATTTATCAAGCTCTCAACTAT 435				
OY	279 CATTCTACATCACTCAATATATCTTACGCTAAA-TGTTGAACAT-AAACCATAG 336				
DB	436 CATTCTACATCACTCAATATATCTTACGCTAAA-TGTTGAACAT-AAACCATAG 495				
OY	337 TCAGCTTGATAGATTAAGCCGTGACATGACAAATACAGACTACAGCAAGATTAAG 396				
DB	496 TCATATACGATGATCAATTTGGGTGTGAACATGCAACCAACGAGAGTCAAGCAAGATTAAG 555				









	RESULT 4
ID	AAL38432
XX	AAL38432 standard; DNA; 4505 BP.
XX	
AC	AAL38432;
DE	15-AUG-2002 (first entry)
DE	
XX	Plasmid pAP7249 encoding secoisolaricresinol dehydrogenase.
KW	Guaiaacyl (g)-ligan; monooctyledon plant; rice; food additive;
XX	seed-specific transcriptional regulatory region; dehydrogenase; gene; ds.
OS	Unidentified.
FT	
FT	Location/Qualifiers
FT	CDS
FT	/tag= a
FT	/product= "Protein of secoisolaricresinol dehydrogenase"
FT	/tag= b
FT	/product= "Protein of Kanamycin"
PN	WO200220548-A1.
XX	
PD	14-MAR-2002.
PF	
PF	04-SEP-2001; 2001WO-US27500.
PR	
PR	07-SEP-2000; 2000US-230632P.
PA	(UNIT ) UNIV WASHINGTON STATE RES FOUND;
PA	(PHYT-) APPLIED PHYTOLOGICS INC.
PI	
PI	Lewis NG, Davin LB, Huang N;
PS	WPI; 2002-425767/45.
DR	P-PSDB; AAO21494, AAO21498.
XX	
PT	Increasing guaiaacyl-lignan content in monocotyledon plants, by
PT	transforming plant with chimeric gene construct having seed-specific
PT	transcriptional regulator linked to gene encoding protein involved in
PT	G-lignan formation -
XX	
PS	Claim 9; Fig 10; 136pp; English.
XX	
CC	The invention relates to a method for increasing the guaiaacyl (g)-ligan
CC	content in seeds of a monocotyledon plant, comprising selecting at least
CC	one protein or enzyme integral to the pathway leading to G-lignan
CC	formation, stably transforming a monocotyledon plant with chimeric gene
CC	(CG) constructs having a seed-specific transcriptional regulatory region
CC	operably linked to a nucleic acid sequence encoding the enzyme. The
CC	method of the invention is useful for stably transforming a
CC	monocotyledonous plant (e.g. rice) with CG constructs resulting in
CC	increased expression of the genes encoded by CG constructs. The G-lignan
CC	enchilised seed composition is useful as a food additive. This
CC	polynucleotide sequence represents plasmid PAP7249 encoding
CC	secoisolaricresinol dehydrogenase relating to the invention.
XX	
SQ	Sequence 4505 BP; 1166 A; 1096 C; 1099 G; 1144 T; 0 other;
	Query Match 90.8%; Score 707; DB 24; Length 4505;
	Best local similarity 99.3%; Pred No. 3.2e-179;
	Matches 710; Conservative 0; Mismatches 5; Indels 0; Gaps 0
OY	
	65 TCATGAGTATGTCTGAACAATTATGGACACGCAATAAAGAACAATTTGATGAGTCG 124
Db	238 TCATGAGTATGTCTGAACAATTATGGACACGCAATAAAGAACAATTTGATGAGTCG 297
OY	
	125 TGTATCTTGATGAGACCTCCTAAAGTTCTCTACCCCGGTAAGAAACCCTTAAGCAAATG 184
Db	298 TGTATCTTGATGAGACCTCCTAAAGTTCTCTACCCCGGTAAGAAACCCTTAAGCAAATG 357

QY	188	GC	AAAGTTGGATTTC	CACTGACATTAATGCAAAATAGATATCATCATGATGACATAGCA	244
Db	358	GC	AAAGTTGGATTTC	CACTGACATTAATGCAAAATAGATATCATCATGATGACATAGCA	417
QY	245	CT	CATGATCATATCA	AGCCTCTCTCAACCTATTCATTTCTTCTACTCATCTACATAGATC	304
Db	418	CT	CATGATCATATCA	AGCCTCTCTCAACCTATTCATTTCTTCTACTCATCTACATAGATC	477
QY	305	TT	CAGCTTAATGTTAGA	CACTAAACCATTAAGTCACTTTGANGAGTATYBGGGTGACA	364
Db	478	TT	CAGCTTAATGTTAGA	CACTAAACCATTAAGTCACTTTGANGAGTATYBGGGTGACA	537
QY	365	CAT	GCAATACAG	AGACTAGCAAAATGATATGATCAATAAACATCACA	424
Db	538	CAT	GCAATACAG	AGACTAGCAAAATGATATGATCAATAAACATCACA	597
QY	425	GCT	ATATGTCATAT	TGCAAAAAGAGAGGCTTATTAAGACAAAGCATGACTCACAAAAT	484
Db	598	GCT	ATATGTCATAT	TGCAAAAAGAGAGGCTTATTAAGACAAAGCATGACTCACAAAAT	657
QY	485	TC	ACTGGCTTTGGTGT	GCACAAAAGAGAGGCTTTAATATTCANGTCAATTTGCCAAA	544
Db	658	TC	ACTGGCTTTGGTGT	GCACAAAAGAGAGGCTTTAATATTCANGTCAATTTGCCAAA	711
QY	545	GAA	AGAGAAAGACAA	CAACATCTCGTCATATTAACATATTCGATGTCATCAAT	604
Db	718	GAA	AGAGAAAGACAA	CAACATCTCGTCATATTAACATATTCGATGTCATCAAT	777
QY	605	AT	CTACCACTTCTGT	TACACACTTCATATATCAATTAAGATCACCTTCACGCTGGAC	664
Db	778	AT	CTACCACTTCTGT	TACACACTTCATATATCAATTAAGATCACCTTCACGCTGGAC	837
QY	665	ATT	AACAACACATCT	CTAATCTTAATGGAAGAGCTTTATCTACATATTAAGACAG	724
Db	838	ATT	AACAACACATCT	CTAATCTTAATGGAAGAGCTTTATCTACATATTAAGACAG	897
QY	725	ANG	ATTTCTATGTTTCT	CAAAAAAGCAATTCATTACTTACTTACAAAC	779
Db	898	ANG	ATTTCTATGTTTCT	CAAAAAAGCAATTCATTACTTACTTACAAAC	952
RESULT 5					
AAD45308					
ID	AAD45308	standard; DNA; 786 BP.			
XX	AC	AAD45308;			
XX	AC	AAD45308;			
XX	27-DEC-2002	(first entry)			
XX	DE	Rice Gtl (glutelin) promoter-Gtl leader DNA.			
XX	DE	Rice Gtl (glutelin) promoter-Gtl leader DNA.			
XX	KW	Rice; feed improvement; plant-derived feed; antibiotic; additive;			
XX	KW	anti-microbial; poultry; alpha-1-antitrypsin; flour; malt; promoter;			
XX	XX	glutelin; Gtl; ds.			
XX	OS	Oryza sativa.			
XX	EN	WO200263975-A2.			
XX	22-AUG-2002.				
XX	PD				
XX	PF	14-FEB-2002; 2002WO-US04919.			
XX	14-FEB-2001; 2001US-269186P.				
XX	FR	02-MAY-2001; 2001US-0847232.			
XX	PA	(VENT-) VENTRIA BIOSCIENCE.			
XX	EL	Huang N, Rodriguez RL, Hagle FE;			
XX	FR	WPI; 2002-682708/73.			

PT Improved feed for production animals, comprising plant-derived feed  
PT ingredients, and seed composition containing flour, extract, or malt  
PT from mature monocot seeds and heterologous seed-produced anti-microbial  
PT proteins

PS Claim 19; Page 153; 175pp; English.

CC The invention relates to improved feed for production animals,  
CC comprising one or more plant-derived feed ingredients, substantially  
CC unsupplemented with small-molecule antibiotics and as an additive a  
CC seed composition containing a flour, extract or malt obtained from  
CC mature monocot seeds and one or more heterologous seed-produced anti-  
CC microbial proteins in substantially unpurified form. The invention  
CC is useful as a feed for production animals such as poultry and hoofed  
CC farm animals. The present sequence is rice Gtl (glutelin) promoter  
CC and Gtl leader DNA. This sequence is used in the invention.

XX  
XX  
XX Sequence 786 BP; 264 A; 178 C; 124 G; 220 T; 0 other;

Query Match 90.6%; Score 706; DB 24; Length 786;  
Best Local Similarity 99.3%; Pred. No. 3.4e-179;  
Matches 709; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 66 CATGAGTAATGTTGAGCATTTATGAGGACGAAATATAAAGAACTTTGATGAGTCT 125  
DB 1 CATGAGTAATGTTGAGCATTTATGAGGACGAAATATAAAGAACTTTGATGAGTCT 60  
QY 126 GATACCTCGATGAGCCTCAAAAGTTCTCAACCCGGATAAGAAACCTTTAAGCAATGTG 185  
DB 61 GATACCTCGATGAGCCTCAAAAGTTCTCAACCCGGATAAGAAACCTTTAAGCAATGTG 120  
QY 186 CAAAGTTTGCATTTCTCCACTGACATATGCAAAATAGATATCATGATGACATGCAAC 245  
DB 121 CAAAGTTTGCATTTCTCCACTGACATATGCAAAATAGATATCATGATGACATGCAAC 180  
QY 246 TCATGATCATATATGATCCCTCTCAACCTATATTCATCTCAACATCATATATATCT 305  
DB 181 TCATGATCATATATGATCCCTCTCAACCTATATTCATCTCAACATCATATATATCT 240  
QY 306 TCAGCTAAATGTTGAACATTAACCCATTAAGTCAAGCTTTGATGATTAAGGCGTAGAC 365  
DB 241 TCAGCTAAATGTTGAACATTAACCCATTAAGTCAAGCTTTGATGATTAAGGCGTAGAC 300  
QY 366 ATGACAAATACAGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 425  
DB 301 ATGACAAATACAGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 360  
QY 426 CTATATGTCATATTTGCAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 485  
DB 361 CTATATGTCATATTTGCAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420  
QY 486 CACTTGCTTTCTGTGCAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 545  
DB 421 CACTTGCTTTCTGTGCAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480  
QY 546 AAAG 605  
DB 481 AAAG 540  
QY 606 TTCAATCCACTTTCTGTACCACTTCATATATCATATAGAGTCACTCACTGAGACA 665  
DB 541 TTCAATCCACTTTCTGTACCACTTCATATATCATATAGAGTCACTCACTGAGACA 600  
QY 666 TTACAAACTCTATCTTACATTTAGATGCAAGAGAGAGAGAGAGAGAGAGAGAGAG 725  
DB 601 TTACAAACTCTATCTTACATTTAGATGCAAGAGAGAGAGAGAGAGAGAGAGAGAG 660  
QY 726 TGATTTTCATTTGTTCTCAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 779  
DB 661 TGATTTTCATTTGTTCTCAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 714

RESULT 6

AAAD4933  
ID AAAD4933 standard; DNA; 786 BP.

AC AAAD4933;

DF 13-DEC-2002 (first entry)

DE Rice Gtl (glutelin) promoter and Gtl leader DNA.

KW Rice; plant-derived food; flour; malt; monocot seed; milk protein;

KW infant formula; ds.

OS Oryza sativa.

PN WO200264814-A2.

PD 22-AUG-2002.

PF 14-FEB-2002; 2002MO-US04921.

PR 14-FEB-2001; 2001US-269199P.

PR 02-MAY-2001; 2001US-0847232.

P2 (VENT-) VENTRIA BIOSCIENCE.

P1 Huang N, Rodriguez RL, Hagie FE;

DR WPI: 2002-667011/71.

PT New nutritionally enhanced food compositions, useful for improving  
PT infant formula, or supplementing or enhancing the diet of infants,

PS particularly very-low birth weight infants

XX Example 9; Page 107; 179pp; English.

CC The invention relates to a nutritionally enhanced food comprising one  
CC or more plant-derived food ingredients, and as an additive, a seed  
CC composition containing a flour, extract, or malt obtained from mature  
CC monocot seeds and one or more seed-produced human milk proteins in  
CC substantially unpurified form. The nutritionally enhanced food and food  
CC additive compositions are useful for improving infant formula. Infant  
CC formulas containing the recombinant human milk proteins are useful in  
CC supplementing or enhancing the diet of infants, particularly very-low  
CC birth weight infants. The present sequence is rice Gtl (glutelin)  
CC promoter and Gtl leader DNA.

SQ Sequence 786 BP; 264 A; 178 C; 124 G; 220 T; 0 other;

Query Match 90.6%; Score 706; DB 24; Length 786;  
Best Local Similarity 99.3%; Pred. No. 3.4e-179;  
Matches 709; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 66 CATGAGTAATGTTGAGCATTTATGAGGACGAAATATAAAGAACTTTGATGAGTCT 125  
DB 1 CATGAGTAATGTTGAGCATTTATGAGGACGAAATATAAAGAACTTTGATGAGTCT 60  
QY 126 GATACCTCGATGAGCCTCAAAAGTTCTCAACCCGGATAAGAAACCTTTAAGCAATGTG 185  
DB 61 GATACCTCGATGAGCCTCAAAAGTTCTCAACCCGGATAAGAAACCTTTAAGCAATGTG 120  
QY 186 CAAAGTTTGCATTTCTCCACTGACATATGCAAAATAGATATCATGATGACATGCAAC 245  
DB 121 CAAAGTTTGCATTTCTCCACTGACATATGCAAAATAGATATCATGATGACATGCAAC 180  
QY 246 TCATGATCATATATGATCCCTCTCAACCTATATTCATCTCAACATCATATATATCT 305  
DB 181 TCATGATCATATATGATCCCTCTCAACCTATATTCATCTCAACATCATATATATCT 240  
QY 306 TCAGCTAAATGTTGAACATTAACCCATTAAGTCAAGCTTTGATGATTAAGGCGTAGAC 365  
DB 241 TCAGCTAAATGTTGAACATTAACCCATTAAGTCAAGCTTTGATGATTAAGGCGTAGAC 300  
QY 366 ATGACAAATACAGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 425



```

Db      301 ATGACAAATTCACAGCTCAAGCAGATTAAGCAAAATGATGTATTAATTAATCCAGAG
Qy      426 CTATATGTCATATTGCAAAAAGAGAGAGAGCTTATTAAGACAGGCACTACACAAAATT
Db      361 CTATATGTCATATTGCAAAAAGAGAGAGCTTATTAAGACAGGCACTACACAAAATT
Qy      486 CACTTGCCTTCCTGCTGTCAAAAAGAGAGGCTTTACATATCCATGCTCATTTGCAAAAG
Db      421 CACTTGCCTTCCTGCTGTCAAAAAGAGAGGCTTTACATATCCATGCTCATTTGCAAAAG
Qy      546 AAAGAGAGAAAGAAACAACAATGCTGGCTCAATTAATCAATCTGATGTCACATTA
Db      481 AAAGAGAGAAAGAAACAACAATGCTGGCTCAATTAATCAATCTGATGTCACATTA
Qy      606 TTCAATCCACCTTTGCTGTACCAACCTTCATATATCAATTAAGAGTCACTTCACTGAC
Db      541 TTCAATCCACCTTTGCTGTACCAACCTTCATATATCAATTAAGAGTCACTTCACTGAC
Qy      666 TTAACAAACCTATCTTAATCAATTTAAGAGAGAGGCTTTATCTCATCTATAATGCAAGA
Db      601 TTAACAAACCTATCTTAATCAATTTAAGAGAGAGGCTTTATCTCATCTATAATGCAAGA
Qy      726 TGATTTCTCATTTGTTCTCCAAAAGAGATTCAGTTCAATAGTCTCAACAAC
Db      661 TGATTTCTCATTTGTTCTCCAAAAGAGGCGGCTTCAATAGTCTCAACAAC 714

```

RESULT 7  
 ABS66526  
 ID ABS66526 standard; DNA; 786 BP.

AC ABS66526;

DT 29-NOV-2002 (first entry)

DE Rice Glutelin (Gt) 1 promoter and Gt1 leader sequence.

KM Transformed plant; heterologous transcription factor; transgenic plant;  
 KM seed protein; protein expression; gene expression control sequence;

KW ss.

OS Oryza sativa.

PN W0200264750-A2.

PD 22-AUG-2002.

PF 14-FEB-2002; 2002MO-US04909.

PR 14-FEB-2001; 2001US-269188P.

PR 14-FEB-2001; 2001US-269199P.

PR 02-MAY-2001; 2001US-0847232.

PA (VENT-) VENTRIA BIOSCIENCE.

PI Huang N, Yang D;

DR WPI: 2002-657592/70.

PT Producing heterologous polypeptide in plant grain, by culturing  
 PT transformed plant to form a grain-producing transforming plant, and  
 PT recovering transgenic grains containing polypeptide from transformed  
 PT plant

XX Example 15; Page 119-120; 230pp; English.

CC The invention describes a method of producing a heterologous polypeptide  
 CC (1) in a grain of a plant, comprising culturing a heterologous polypeptide  
 CC comprising a first chimeric gene, and optionally, at least one  
 CC heterologous transcription factor that is capable of enhancing the  
 CC expression of the chimeric gene, to form a grain producing transforming  
 CC plant (P2), and recovering transgenic grains containing (1) from P2.

CC The method is useful for producing heterologous polypeptide in a grain  
 CC of a plant. (1) is a non-plant storage, human or non-human animal, milk  
 CC or other than a milk polypeptide, antibodies, cytokines, lymphokines,  
 CC chemokines, hormones, growth factors, coagulation factors,  
 CC anti-infectives, or cytotoxins, or anti-inflammatory molecule or  
 CC intestinal trefoil factor (ITF) or its active fragment. Preferably, (1)  
 CC is lactoferrin, lysozyme, lactoferricin, ITF, epidermal growth factor  
 CC (EGF), keratinocyte growth factor (KGF), insulin-like growth factor I  
 CC (IGF-I), lactohedinin, kappa-casein, naplocorin, lactoperoxidase,  
 CC alpha-1-antitrypsin, immunoglobulin, alpha-lactalbumin,  
 CC beta-lactoglobulin, alpha-casein, beta-casein, albumin, fibrinogen or  
 CC protease inhibitor. This sequence represents a gene expression control  
 CC sequence used in the method of producing a transgenic plant.

XX Sequence 786 BP; 264 A; 178 C; 124 G; 220 T; 0 other;

Query Match 90.6%; Score 706; DB 24; Length 786;  
 Best Local Similarity 99.3%; Pred. No. 3.4e-179;  
 Matches 709; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

Qy      66 CATGAGTATGTGTGACATTATGAGACACGAAATTAAGAAACAATTTGATGATGCT 125
Db      1 CATGAGTATGTGTGACATTATGAGACACGAAATTAAGAAACAATTTGATGATGCT 60
Qy      126 GTATCCTGATGAGCCTCAAAAGTCTCCACCCCGGATTAAGAAACCTTAAGCATGTG 185
Db      61 GTATCCTGATGAGCCTCAAAAGTCTCCACCCCGGATTAAGAAACCTTAAGCATGTG 120
Qy      186 CAAAGTTGCATTCCTCAGCATGACATTAATGCAAAATTAAGATATCATGACATGCAAC 245
Db      121 CAAAGTTGCATTCCTCAGCATGACATTAATGCAAAATTAAGATATCATGACATGCAAC 180
Qy      246 TCAATGCATATCATGATGCTCTCTCAACCTATTCATCTCATCATCATATATATCT 305
Db      181 TCAATGCATATCATGATGCTCTCTCAACCTATTCATCTCATCATCATATATATCT 240
Qy      306 TCAATGCATATGTTGAACATTAACCCATTAAGAGAGGCTTTGATAGATTAAGAGAGAC 365
Db      241 TCAATGCATATGTTGAACATTAACCCATTAAGAGAGGCTTTGATAGATTAAGAGAGAC 300
Qy      366 ATGACAAATTCACAGCTCAAGCAAGTAAAGCAAAATGATGATCAATTAATCTCCAGAG 425
Db      301 ATGACAAATTCACAGCTCAAGCAAGTAAAGCAAAATGATGATCAATTAATCTCCAGAG 360
Qy      426 CTATATGTCATATTGCAAAAAGAGAGAGCTTATTAAGACAGGCACTACACAAAATT 485
Db      361 CTATATGTCATATTGCAAAAAGAGAGAGCTTATTAAGACAGGCACTACACAAAATT 420
Qy      486 CACTTGCCTTCCTGCTGTCAAAAAGAGAGGCTTTACATATCCATGCTCATTTGCAAAAG 545
Db      421 CACTTGCCTTCCTGCTGTCAAAAAGAGAGGCTTTACATATCCATGCTCATTTGCAAAAG 480
Qy      546 AAAGAGAGAAAGAAACAACAATGCTGGCTCAATTAATCAATCTGATGTCACATTA 605
Db      481 AAAGAGAGAAAGAAACAACAATGCTGGCTCAATTAATCAATCTGATGTCACATTA 540
Qy      606 TTCAATCCACCTTTGCTGTACCAACCTTCATATATCAATTAAGAGTCACTTCACTGAC 665
Db      541 TTCAATCCACCTTTGCTGTACCAACCTTCATATATCAATTAAGAGTCACTTCACTGAC 600
Qy      666 TTAACAAACCTATCTTAATCAATTTAAGAGAGAGGCTTTATCTCATCTATAATGCAAGA 725
Db      601 TTAACAAACCTATCTTAATCAATTTAAGAGAGAGGCTTTATCTCATCTATAATGCAAGA 660
Qy      726 TGATTTCTCATTTGTTCTCCAAAAGAGATTCAGTTCAATAGTCTCAACAAC 779
Db      661 TGATTTCTCATTTGTTCTCCAAAAGAGGCGGCTTCAATAGTCTCAACAAC 714

```

RESULT 8  
 AAL38426  
 ID AAL38426 standard; DNA; 4227 BP.

XX

AC AAL38426;  
 XX  
 DC 15-AUG-2002 (first entry)  
 XX  
 DE Plasmid pAB1244 encoding dirigent protein.  
 XX  
 KW Guaiacyl (G)-ligan: monooctyledon plant; rice; food additive;  
 KW seed-specific transcriptional regulatory region; dirigent; gene; ds.  
 OS  
 XX  
 FH Unidentified.  
 FT  
 FT Key Location/Qualifiers  
 FT CDS 953..1513  
 FT /\*tag= a  
 FT /product= "Protein of dirigent gene"  
 FT CDS 2533..3327  
 FT /\*tag= b  
 FT /product= "Protein of kanamycin"  
 PN W0200220548-A1.  
 XX  
 PD 14-MAR-2002.  
 XX  
 PE 04-SEP-2001; 2001WO-US27500.  
 XX  
 PR 07-SEP-2000; 2000US-230632P.  
 XX  
 PA (UNITN ) UNIV WASHINGTON STATE RES FOUND.  
 PA (PHYT-) APPLIED PHYTOLOGICS INC.  
 XX  
 PI Lewis NG, Davin LB, Huang N;  
 XX  
 DR WPI; 2002-425767/45.  
 DR P-PSDB; AAO21492; AAO21496.  
 XX  
 XX  
 PT Increasing guaiacyl-lignan content in monooctyledon plants, by  
 PT transforming plant with chimeric gene construct having seed-specific  
 PT transcriptional regulator linked to gene encoding protein involved in  
 PT G-lignan formation -  
 XX  
 XX  
 PS Claim 9; Fig 8; 136pp; English.  
 XX  
 CC The invention relates to a method for increasing the guaiacyl (G)-ligan  
 CC content in seeds of a monooctyledon plant, comprising selecting at least  
 CC one protein or enzyme integral to the pathway leading to G-lignan  
 CC formation, stably transforming a monooctyledon plant with chimeric gene  
 CC (CG) constructs having a seed-specific transcriptional regulatory region  
 CC operably linked to a nucleic acid sequence encoding the enzyme. The  
 CC method of the invention is useful for stably transforming a  
 CC monooctyledonous plant (e.g. rice) with CG constructs resulting in  
 CC increased expression of the genes encoded by CG constructs. The G-lignan  
 CC enriched seed composition is useful as a food additive. This  
 CC polynucleotide sequence represents plasmid pAB1244 encoding the dirigent  
 CC protein gene relating to the invention.  
 CC  
 XX  
 SQ Sequence 4227 BP; 1066 A; 1058 C; 1026 G; 1077 T; 0 other;

Query Match 88.4%; Score 689; DB 24; Length 4227;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-174;  
 Matches 689; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 65 TCATGAGTAATGTGTGAGCATTAATGGACCAAGAAATAAAGAAATTTTATGAGATCG 124  
 |||  
 DB 238 TCATGAGTAATGTGTGAGCATTAATGGACCAAGAAATAAAGAAATTTTATGAGATCG 297  
 |||  
 QY 125 TGTATCTCGATGAGGCTCAAAAGTTCTCTACCCCGAATGAAGAAACCTTAAGCAATGT 184  
 |||  
 DB 298 TGTATCTCGATGAGGCTCAAAAGTTCTCTACCCCGAATGAAGAAACCTTAAGCAATGT 357  
 |||  
 QY 185 GCAAGTTTGCATTTCCCTGACATTAATGCAAAATAGATATCATGATGACATAGCAA 244  
 |||  
 DB 358 GCAAGTTTGCATTTCCCTGACATTAATGCAAAATAGATATCATGATGACATAGCAA 417  
 |||

QY 245 CTCATGCAATCATATCATGCGCTCTCTCAACGCTATTCCTTACTCATCTACATAGTATC 304  
 |||  
 DB 418 CTCATGCAATCATATCATGCGCTCTCTCAACGCTATTCCTTACTCATCTACATAGTATC 477  
 |||  
 QY 305 TTCAGCTTAATGTTAGACATAAACCCATTAAGTACGTTTGAAGATATTAAGCGGTGACA 364  
 |||  
 DB 478 TTCAGCTTAATGTTAGACATAAACCCATTAAGTACGTTTGAAGATATTAAGCGGTGACA 537  
 |||  
 QY 365 CATGACAAATCACAGACTCAGACAGATTAAGCAAAATGATGTGATCAATTAATTCACAGA 424  
 |||  
 DB 538 CATGACAAATCACAGACTCAGACAGATTAAGCAAAATGATGTGATCAATTAATTCACAGA 597  
 |||  
 QY 425 GCTATATGTCATATGGAAG 484  
 |||  
 DB 598 GCTATATGTCATATGGAAG 657  
 |||  
 QY 485 TCATCTGCGCTTCTGTCGCAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 544  
 |||  
 DB 658 TCATCTGCGCTTCTGTCGCAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 717  
 |||  
 QY 545 GAAAG 604  
 |||  
 DB 718 GAAAG 777  
 |||  
 QY 605 ATTCATCCACCTTCGCTGATACCACTTCATATATATATATATATATATATATATATATATAT 664  
 |||  
 DB 778 ATTCATCCACCTTCGCTGATACCACTTCATATATATATATATATATATATATATATATATAT 837  
 |||  
 QY 665 ATTACAAACCTATCTATTAACATTTAGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 724  
 |||  
 DB 838 ATTACAAACCTATCTATTAACATTTAGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 897  
 |||  
 QY 725 ATGATTTCTCATTTGTTTCTCAAAAAAGC 753  
 |||  
 DB 898 ATGATTTCTCATTTGTTTCTCAAAAAAGC 926  
 |||

RESULT 9  
 AAL38429  
 ID AAL38429 standard; DNA; 4605 BP.  
 XX  
 AC AAL38429;  
 XX  
 DC 15-AUG-2002 (first entry)  
 XX  
 DE Plasmid pAB1246 encoding pinorensinol/lariciresinol reductase.  
 XX  
 KW Guaiacyl (G)-ligan: monooctyledon plant; rice; food additive;  
 KW seed-specific transcriptional regulatory region; reductase; gene; ds.  
 OS  
 XX  
 FH Unidentified.  
 FT  
 FT Key Location/Qualifiers  
 FT CDS 953..1891  
 FT /\*tag= a  
 FT /product= "Protein of pinorensinol/lariciresinol  
 FT reductase"  
 FT CDS 2911..3705  
 FT /\*tag= b  
 FT /product= "Protein of kanamycin"  
 PN W0200220548-A1.  
 XX  
 PD 14-MAR-2002.  
 XX  
 PE 04-SEP-2001; 2001WO-US27500.  
 XX  
 PR 07-SEP-2000; 2000US-230632P.  
 XX  
 PA (UNITN ) UNIV WASHINGTON STATE RES FOUND.  
 PA (PHYT-) APPLIED PHYTOLOGICS INC.  
 XX  
 PI Lewis NG, Davin LB, Huang N;



Db	238	TCATGAGTAATGTGTGGAGCACTTATGGGACACAGAAAATAAAGAAACATTTTGTATGATGCG	297
QY	125	TGATATCCGTGATGAGGCGCTCAAAAGTTCTCTCCACCCCGGATAGAAACCGCTTAAGCAATGT	184
Db	238	TGATATCCGTGATGAGGCGCTCAAAAGTTCTCTCCACCCCGGATAGAAACCGCTTAAGCAATGT	357
QY	135	GCAAAAGTTTGCATTTCCTCCATGACATTAATGCAAAAAPAGATATCATGATGACATAGCAA	244
Db	358	GCAAAAGTTTGCATTTCCTCCATGACATTAATGCAAAAAPAGATATCATGATGACATAGCAA	417
QY	245	CTCATGCGATCAATCAATGAGCGCTGCTCCACACATTCATCTCCATGCTCATGCTCAATGATC	304
Db	418	CTCATGCGATCAATCAATGAGCGCTGCTCCACACATTCATCTCCATGCTCATGCTCAATGATC	477
QY	305	TTCAAGCTAAATGTTAGAACATAAACCATAAGTCAGTTTGAATGATTAAGGCGTGACA	364
Db	478	TTCAAGCTAAATGTTAGAACATAAACCATAAGTCAGTTTGAATGATTAAGGCGTGACA	537
QY	365	CATGACAATATCAACAGACTCAGCAGATAGAAAGCAAAATAGTGTACATATAACCTCCAGA	424
Db	538	CATGACAATATCAACAGACTCAGCAGATAGAAAGCAAAATAGTGTACATATAACCTCCAGA	597
QY	425	GCTATATGTCAATATTGCAAAAAGAGAGAGTTTAAAGACAAAGGCGATGACTCCAAAAAT	484
Db	598	GCTATATGTCAATATTGCAAAAAGAGAGAGTTTAAAGACAAAGGCGATGACTCCAAAAAT	657
QY	455	TCGATGCGCTTTCGCTGTCCAAAAGGAGGAGCGCTTTCATATTATCCATGTCAATTTGCAAA	544
Db	658	TCGATGCGCTTTCGCTGTCCAAAAGGAGGAGCGCTTTCATATTATCCATGTCAATTTGCAAA	717
QY	545	GAAAGAGAGAAAGAAACACACATAGCTGCGTCATTTTACATATCTGATGTCATCATATT	604
Db	718	GAAAGAGAGAAAGAAACACACATAGCTGCGTCATTTTACATATCTGATGTCATCATATT	777
QY	605	ATTCATCCACGCTTGTGTGTACACACACTTCATATATATCATATAAGAGTCATTCACGCTGTGAC	664
Db	778	ATTCATCCACGCTTGTGTGTACACACACTTCATATATATCATATAAGAGTCATTCACGCTGTGAC	837
QY	665	ATTATACAACTCTATCTTAACATTTTGAATGGAAGAGCGCTTATTCACGCTTAATAGTCAAG	724
Db	838	ATTATACAACTCTATCTTAACATTTTGAATGGAAGAGCGCTTATTCACGCTTAATAGTCAAG	897
QY	725	ATGATTTCTCATTTGTTTCTCACAAAAGGATTCAGATT	761
Db	898	ATGATTTCTCATTTGTTTCTCACAAAAGGAGCGCGCTT	934
RESULT 11			
AAL38419			
ID AAL38419 standard; DNA; 690 BP.			
AAL38419;			
XX	15-AUG-2002	(first entry)	
DT			
XX			
DE	Gt-1 promoter polynucleotide sequence.		
XX			
KW	Guaiacyl (G)-lignan; monocotyledon plant; rice; food additive;		
XX	seed-specific transcriptional regulatory region; Gt 1 promoter; ds.		
OS	Unidentified.		
XX			
PN	MO200220548-A1.		
XX			
PD	14-MAR-2002.		
XX			
PE	04-SEP-2001; 2001MO-US27500.		
XX			
PR	07-SEP-2000; 2000US-230632P.		
XX			
PA	(UNIV ) UNIV WASHINGTON STATE RES FOUND.		
XX	PA (PHYT-) APPLIED PHYTOLOGICS INC.		

Query Match	Best Local Similarity	88.3%	Score 688;	DB 24;	Length 690;
Matches 688;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0	
66	CATGATGATGTGAGACATTTATGAGCACCAGAAATATAAGACATTTGATGATGCT	125			
1	CATGATGATGTGAGACATTTATGAGCACCAGAAATATAAGACATTTGATGATGCT	60			
126	GATGCTTCGATGAGCCTCAAAAGTCTCTACCCCGATTAAGAAACCTTAAAGCATGTG	185			
61	GATGCTTCGATGAGCCTCAAAAGTCTCTACCCCGATTAAGAAACCTTAAAGCATGTG	120			
186	CAAACTTTGACATTCCTCCACGACATTAAGCAAAATTAAGATATCATGATGATGACATGAC	245			
121	CAAAGTTCGATTCCTCCACGACATTAAGCAAAATTAAGATATCATGATGATGACATGAC	180			
246	TCATGATCATATCATGATGATGCTCTCTCAACCTATTCATTCCTCATCTPACATTAAGATCT	305			
181	TCATGATCATATCATGATGATGCTCTCTCAACCTATTCATTCCTCATCTPACATTAAGATCT	240			
306	TCAGCTAAATGTTGAAATCAATAAACCATAATCAACCTGTTATGATGATTAAGCGCTGAC	365			
241	TCAGCTAAATGTTGAAATCAATAAACCATAATCAACCTGTTATGATGATTAAGCGCTGAC	300			
366	ATGACAAATTCACACAGATCAAGCAAGATTAAGCAAAATGATGTATACATTAATCTCCAG	425			
301	ATGACAAATTCACACAGATCAAGCAAGATTAAGCAAAATGATGTATACATTAATCTCCAG	360			
426	CTATATGTCATTTGCAAAAAAGAGAGAGCTTTATTAAGACAAGGATGATGATCAAAAAAT	485			
361	CTATATGTCATTTGCAAAAAAGAGAGAGCTTTATTAAGACAAGGATGATGATCAAAAAAT	420			
486	CATTGCTCTTCCTGTGTAAATAAGAGAGAGGCTTACATTTCCATGATGATGATGATGAA	545			
421	CATTGCTCTTCCTGTGTAAATAAGAGAGAGGCTTACATTTCCATGATGATGATGATGAA	480			
546	AAAGAGAGAAAGAAACAACAATGCTGCTCATTTATTAATCAATTCGTATGTCATCATTA	605			
481	AAAGAGAGAAAGAAACAACAATGCTGCTCATTTATTAATCAATTCGTATGTCATCATTA	540			
606	TTATGCTGCTGCTGCTGATGACACATTCATTAATCAATTAAGATGATGATGATGATGAA	665			
541	TTATGCTGCTGCTGCTGATGACACATTCATTAATCAATTAAGATGATGATGATGATGAA	600			
666	TTATGCTGCTGCTGCTGATGACACATTCATTAATCAATTAAGATGATGATGATGATGAA	725			
601	TTATGCTGCTGCTGCTGATGACACATTCATTAATCAATTAAGATGATGATGATGATGAA	660			





CC of the printed specification, but was obtained in electronic  
CC format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
xx

SQ Sequence 61020 BP; 17884 A; 634 C; 12537 G; 29965 T; 0 other;

Query Match 6.2%; Score 48; DB 22; Length 61020;

Best Local Similarity 46.8%; Pred. No. 0.023;

Matches 220; Conservative 0; Mismatches 245; Indels 5; Gaps 2;

```
QY 14 ACAAACTBAAGTAAGAGAGAGTGGGTGTTAGAAAAGAAACAAATATCATGAGTA 73
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 54905 AAAATCTAATTAATAAATTAACCTAATTAATTAATTAACAAAAAATAATCA 54846

QY 74 AAGTGAAGCAATTAAGGACCAAGAAATTAATAAGACATTTTGATGAGTGGTATCCTC 133
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 54845 AGGATTAACCTCTATATCTAATAACTTAATAAATCTAATAAATTTACCTACATCCCC 54786

QY 134 GATGAGCCTCAAAAGTTCTCTCAACCCCGATPAAGAAACCTTAAGCAATGTGCAAGTTT 193
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 54785 GTTACAATCAACAAACCCCTCCGACTTCTATTAATTTTAAAAAATCTAACAAATTT 54726

QY 194 GCATCTCAGCTGACATTAAGCAAAATTAAGATATCAAGATGACATGCACTCATGCAAT 253
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 54725 ACATTAATCCCTCAAAAAAATAAAGCTGCTCGAAACTTACTCATCCTCAATCA 54666

QY 254 CATATCAGCCTCTCTCTCAACCTATTCATTCCTACTCATCTAATTAATCTTCTCACTTA 313
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 54665 ATTACAAAAAATCTCAAAAAAATTTCACTTAAGCACT - ATAAATATATTAATAAAA 54608

QY 314 AAGTTGAACATTAACCCATTAAGTCACTGTTGATGAGTATTAGCGTGAACATGACAAA 373
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 54607 AATTAATTAACACTTCTCTTAACCAAAATTCCTTCTTCCCATTTCAACGAAAAA 54548

QY 374 TCAAGAGCTCAGCAAGATTAAGCAAAATGATGTACATPAACATCCAGAGCTATATGT 433
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 54547 ACCCCCAAAAAATCTCTCCCT--AAATTAACCTCTCCATTAACCTCCCTTAATTAATA 54491

QY 434 CATATTGCAAAAAGAGAGAGCTTAATPAAGCAAGGCATGACTCAAAAAA 483
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 54490 AATATTAAAAAACAACAAAAAATAAATAAATAAATAAATAAATAAATAAATAA 54441
```

Search completed: September 13, 2003, 19:02:12  
Job time : 134.962 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 13, 2003, 02:54:05 ; Search time 1025.62 Seconds  
(without alignments)  
18460.230 Million cell updates/sec

Title: US-09-847-232a-26

Perfect score: 1 ttcctgagtaagacaaac.....ttcattagtcctacaaac 779

Sequence:

IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 1215238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :  
1: em\_estba:\*  
2: em\_estbun:\*  
3: em\_estlin:\*  
4: em\_estlu:\*  
5: em\_estlv:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_estl:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pin:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_man:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rtd:\*  
26: em\_gss\_png:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	ID	Description
C 1	62	8.0	1180	13	BX436369 BX436369
C 2	58.6	7.5	1124	13	BX436282 BX436282
C 3	53	6.8	712	13	BX416727 BX416727
C 4	49.4	6.3	1101	29	CNS00067 AL062049 Drosophila

Result No.	Score	Match	Length	ID	Description
C 5	48.8	6.3	1101	29	CNS00LT2 AL078714 Drosophila
C 6	48.6	6.2	1031	29	CNS00CF2 AL059199 Drosophila
C 7	48.2	6.2	1091	13	BX424950 BX424950
C 8	48.2	6.2	1201	9	AL567002 BX424950
C 9	48.8	6.2	1101	29	CNS0039G AL567002
C 10	47.8	6.1	804	28	AQ666229 nBdb00270
C 11	47.6	6.1	836	29	B2568401 PUGNHL17B
C 12	47.6	6.1	858	29	B2568398 PUGNHL17B
C 13	47.6	6.1	911	29	CNS00D2S AL060558 Drosophila
C 14	47.4	6.1	802	29	CNS01241 AL101100 Drosophila
C 15	47.7	6.0	563	13	BQ595969 BQ595969
C 16	47.7	6.0	595	13	BQ51276 BQ51276
C 17	46.4	6.0	1101	29	CNS0100X AL098379 Drosophila
C 18	46.2	5.9	680	14	BY17385 BY17385
C 19	46.2	5.9	1137	13	BX444354 BX444354
C 20	46.2	5.9	734	29	B2510892 BQMR139TE
C 21	46.5	5.9	994	13	BX414650 BX414650
C 22	45.8	5.9	759	29	CNS060XV AL411257 T7 end of
C 23	45.8	5.9	1013	14	CD387611 CD387611
C 24	44.8	5.8	1101	29	CNS00HX9 AL073856 Drosophila
C 25	44.6	5.7	970	14	CD051179 AGENCCOURT
C 26	44.6	5.7	1179	29	CNS0607Y AL410324 T7 end of
C 27	44.4	5.6	885	13	BX425603 BX425603
C 28	43.8	5.6	649	12	B1355244 GM30079.5
C 29	43.8	5.6	909	13	BX391298 BX391298
C 30	43.8	5.6	1101	29	CNS017XE AL108152 Drosophila
C 31	43.8	5.6	1111	14	CD389086 CD389086
C 32	43.8	5.6	1201	13	BX406428 BX406428
C 33	43.8	5.6	1201	13	BX444445 BX444445
C 34	43.8	5.6	1201	29	CNS015Y3 AL106053 Drosophila
C 35	43.4	5.6	518	10	BG363885 dg71h1.1
C 36	43.2	5.5	544	12	B1639459 SD22037.3
C 37	43.2	5.5	1201	13	BX459798 BX459798
C 38	43.2	5.5	1052	29	CNS01563 AL105045 Drosophila
C 39	43.2	5.5	1201	13	BX41824 BX41824
C 40	42.8	5.5	983	13	BX404258 BX404258
C 41	42.8	5.5	1058	14	CD048666 AGENCCOURT
C 42	42.8	5.5	1169	29	CNS060XO AL402900 T3 end of
C 43	42.6	5.5	1200	13	BX415897 BX415897
C 44	42.4	5.4	1098	13	BX436460 BX436460
C 45	42.4	5.4	1101	29	CNS00EVL AL069706 Drosophila

## ALIGNMENTS

RESULT 1  
LOCUS BX436369 Homo sapiens THYMUS Homo sapiens CDNA clone C50CAP001YM12  
DEFINITION 1180 bp mRNA linear EST 15-MAY-2003  
5-PRIME mRNA sequence.  
ACCESSION BX436369  
VERSION BX436369.1 GI:30787529  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1180)  
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 Evry cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 516.f for  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=C50CAP001B006PL&cluster=516.f. Contact :  
Feng Liang Email : fliang@life.techn.com URL :  
http://fulllength.invitrogen.com/ invitrogen Corporation 1600



Faraday Avenue Genoscope sequence ID : CS0CAP001BG06QPI  
Location/Qualifiers

[illegible]

	a	c	g	t	others
BASE COUNT	129	29	109	435	422
ORIGIN					



[illegible]

		RESULT	6
CNS00CF2/c			
LOCUS	Drosophila melanogaster genome survey sequence TET3 end of BAC # BACR2SK05 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.		
DEFINITION	BACR2SK05 of RPCI-98 library from Drosophila melanogaster (fruit fly)		
ACCESION	AL059199		
VERSION	AL059199.1 GI:4946662		
KEYWORDS	GSS.		
SOURCE	Drosophila melanogaster (fruit fly)		
ORGANISM	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.		
REFERENCE	1 (bases 1 to 1031) Genoscope. Direct Submission Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr Web : www.genoscope.cns.fr) Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACS. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Ooeegawa and Aaron Mamoser in Pieter de Jong's laboratory in the department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain v3; cn bw sp, the same strain used for the BDGP's pI and EST libraries. A more detailed description of the library and how to order individual BAC clones' the entire library, or filters for hybridization from the BCRC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.		
JOURNAL	location/Oualifiers		
	1..1031		
	/organism="Drosophila melanogaster"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:7227"		
	/clone=BACR2SK05		
	/clone_lib=RPCI-98		
	/note=end:TET3		
	176 a 169 c 124 g 214 t 348 others		
BASE COUNT			
ORIGIN			
Query Match	6.2%; Score 48.6;	DB 29;	Length 1031.
Best Local Similarity	27.7%;	Pred. No. 9.4;	

65 TCATGAGTAATGTGTGAGCATTATGGGACCCACGAAATAAAAAAGAACATTTTCATGAGTCG 12

BASE COUNT	259 a	225 c	203 g	367 t	147 others
vector. Library was not normalized.					

Query Match	6.2%;	Score 48;	DB 29;	Length 1101;
Best Local Similarity	16.8%;	Pred. No. 12;		

Clemson University  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA

Tel: 864 656 7288  
Email: [twinc@clemonson.edu](mailto:twinc@clemonson.edu)  
Fax: 864 656 4293  
Seq primer: TAATACGACTCATCATAGG  
Class: BAC ends  
High quality sequence start: 40  
High quality sequence stop: 212.

## FEATURES

### Source

```

/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/strain="Japonica"
/cultivar="Nipponbare"
/db_xref="caxon:39947"
/clone="hdbp0027020f"
/assay_type="leaf"
/lab_host="E. coli DH10B"
/clone_id="CGI Rice BAC Library (ECORI)"
/notes="Vector: pBACindigo; Site 1: EcorI; Site 2: EcorI;
Rice is the most important food crop in the world. Half of
the world population, especially those inhabiting highly
populated areas of the humid tropics and subtropics, rely
on rice as their primary source of carbohydrate.
Monocotyledonous rice is a diploid plant (2n=24) with a
monocotyledonous genome equivalent of 431 Mbp (Arumuganathan and
Earle, 1991). The relatively small genome of rice, three
times larger than that of Arabidopsis, makes it suitable
for genomic studies. In order to facilitate positional
cloning, physical mapping and genome sequencing of rice,
we have constructed a BAC library from Oryza sativa.
Nipponbare variety using EcoRI as the cloning enzyme. The
library contains 55,296 clones with an average insert size
of 121 kb providing approximately 15 haploid genome
equivalents. The deep coverage allows the isolation a
particular sequence with a probability of 99.9 %. Three
high density filters, each containing 18,432 clones
(doubly spotted), represent the whole library for colony
screening and can be requested from the Clemson University
BAC/EST Resource Center (www.genome.clemson.edu)."
```

[illegible]

RESULT 11

52968401  
10000

LOCUS	836 bp	DNA	linear	GSS 25-MAR-2003
DEFINITION	BUGHNLITD ZM_0.6_1.0_KB Zea mays genomic clone ZMABTa385B2,			

2250

ACCESSION	BZ968401
VERSION	BZ968401.1
	GI:29187123

**KEYWORDS**  
**SOURCE**

ORGANISM:

REFERENCE  
1 (bases 1 to 836)  
cidae; Panicoideae; Andropogoneae; Zea.

**AUTHORS**  
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick,  
A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and Bennettzen, J.

TITLE	Maize Genomics Consortium
JOURNAL	Unpublished
COMMENT	other_gsss: PUGHN1TB

## FEATURES

### Source

```

/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B3"
/db_xref="taxon:4577"
/clone="ZM87A385B2"
/clone_1b="ZM.O.6.1.0_KA"
/notes="vector:PCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
cot selected genomic DNA library"
294 a 155 c 139 g 248 t
BASE COUNT
ORIGIN

```

Query Match	6.1%;	Score 47.6;	DB 29;	Length 836;
Best Local Similarity	49.2%;	Pred. No. 15;		
Matches 125;	Conservative	0;	Mismatches 129;	Indels 0;
				Gaps 0;

OY	309	GCTAAATGTTAGACATAAACCATTGAAGTCACGTTTGATGAGTATTAGGGCGTAGCATG	368
Db	333	GCATATGTGACCAATTAATAATATGATGTGAATAATTAATATTATTAAGTCACAATAAG	392
OY	369	NCAAATCACAGACTCAGCAAGCAATTAAGCCAAATGATGTGTCAATAAACCTCCAGAGCTA	428
Db	393	GAAATTAATAATTAATAAGAAATTCCTGCAATTTTBGAGTGGAAGAAGAAAAGAAATGTG	452
OY	429	TATGCTATTATTGCAAAAAAGSAGAGCTTTAAGACAAGGCAATGCTCACAAAATTCAC	488
Db	455	TTTAAATAATTATTAATTAAGSTAAGAAATTTGTGTGCCAAGCTGATCTCTTAATTGAG	512
OY	489	TTGCGTTTCGTGCATAAAGAGAGAGCGCTTTAOCATTATTCATGTCAATATGCAAAAAGAA	548
Db	513	GAGGATTTTACCAATAGAGGAAATTTGAATTTTCGAACTTTGAATTTCCAACHTCAAAAAAAAA	572
OY	549	GAGAGAAAGAACAA	562
Db	573	GAAAAAACGAAA	586

RESULT 12  
B2968398/c  
LOCUS  
DEFINITION PUGN1ITB\_ZM\_0.6\_1.0\_KB Zee mays genomic clone ZMM87A385B22,  
genomic survey sequence.  
B2968398  
ACCESSION

## RESULT 12

BZ968398

LOCUS

## DEFINITION

genomic survey sequence.  
ACCESSION B2968398

ACCESSION

FEATURES	
source	Location/Qualifiers
1..858	
/organism="Zea mays"	
/mol_type="genomic DNA"	
/strain="B73"	
/db_xref="taxon:4577"	
/clone="ZMMA385B22"	
/clone_lib="ZM 0.6 1.0 kb"	
/note="vector: pCR4-topo. Site1: EcoRI; 0.6-1.0 kb high"	
COT selected genomic DNA library	
200 a 182 c 221 g 235 t	

Query Match	6.1%;	Score 47.6;	DB 29;	Length 858;
Best Local Similarity	49.2%;	Pred. No. 15;		
Matches 125;	Conservative 0;	Mismatches 129;	Indels 0;	Gaps 0;
QY	309	GCTAAATGTTAGAACTAAACCCATTAAGTCAGTTGGATGAGTATTTAGCGGTGACACATG	368	
Db	810	GCAATTTGTACCACTATTAATAATTAATTAATTTGAAATTAATAATTTAAAGTCACAAATG	751	
QY	369	ACAATTCACAGACTCAAGACAAATTAAGCAAAATGATGTCATCAATTAACCTCCAGAGCTA	428	
Db	750	GAAATTTAAATTAATAAAGAAATTTCTGAAATTTAGAGAGGGAAGAAAGAAATTTG	691	
QY	429	TATGTCATTTGGCAAAAAGAGGAGAGCTTTTAAAGACAAGCATGATCCACAAAATTTGAC	488	
Db	690	TATCAAAATTAATAATTAAGSTAAGATATTATGTCACACATGATTTCTCTTAATTTGAG	631	
QY	489	TTGCGCTTTCTGTCCAAAAGAGAGGCGCTTTAACTATTCATGTCATTTGCAAAAGAAA	548	
Db	630	GAGGATTTTACCATTAAGGAGAAATTTGAATTTGCACTTAATTTCAAACTCAAAAAAAA	571	
QY	549	GAGAGAAAGAACAA 562		
Db	570	GAATAAAACAGAAA 557		
RESULT 13				
CNS00D25/c				
LOCUS	CNS00D25	911 bp	DNA	linear
DEFINITION	Drosophila melanogaster genome survey sequence T7 end of BAC #			
	BAC25b07 of RCT-98 library from Drosophila melanogaster (fruit			
	fly), genomic survey sequence.			

RESULT 13				
CNSO0D2S/c				
LOCUS	CNSO0D2S	911 bp	DNA	linear GSS 04-JUN-1996
DEFINITION	<i>Drosophila melanogaster</i> genome survey sequence T7 end of BAC #			
	BAC#2607 of RFLI-98 library from <i>Drosophila melanogaster</i> (fruit fly), genomic survey sequence.			
ACCESSION	AJ060558			
VERSION	AL060558.1	GI:4947516		
KEYWORDS	GSS.			
SOURCE ORGANISM	<i>Drosophila melanogaster</i> (fruit fly)			
	<i>Drosophila melanogaster</i>			
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
	Neoptera; Endopterygota; Diptera; Braconycera; Muscomorpha;			
	Ephyridioidea; Drosophilidae; Drosophila,			

FEATURES	source	Location/Qualifiers
	1. .911	
		/organism="Drosophila melanogaster"
		/mol_type="genomic DNA"
		/db_xref="taxon:7227"
		/clone="BACR26D07"
		/clone_1b="RPCT-98"
		/note="end : T7"
BASE COUNT	140 a	107 c 163 g 251 t 250 others
ORIGIN		

[illegible]

RESULT 14	
CNS01241	
LOCUS	802 bp DNA linear GSS 26-JUL-1999
DEFINITION	CNS01241, <i>Drosophila melanogaster</i> genome survey sequence 77 end of BAC
ACCESSION	BAC07016 of DrosBAC library from <i>Drosophila melanogaster</i> (fruit
VERSION	fly), genomic survey sequence.
KEYWORDS	fly, genomic survey sequence.
SOURCE	AL011100.1 GI:5612711
ORGANISM	GSS.
REFERENCE	<i>Drosophila melanogaster</i> (fruit fly)
AUTHORS	<i>Drosophila melanogaster</i>
TITLE	<i>Drosophila melanogaster</i>
JOURNAL	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
	Neoptera; Endopterygota; Diptera; Brachytera; Muscomorpha;
	Ephydroidea; Drosophilidae; Drosophila.
	1 (bases 1 to 802)
	Genoscope.
	Direct Submission
	Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
	BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
	- Web : www.genoscope.cns.fr)

## COMMENT

Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - <http://www.edgp.ebi.ac.uk> - This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelobAC11.

## FEATURES

Location/Qualifiers  
1..802  
source

/organism="Drosophila melanogaster"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7227"  
/clone\_1lb="BACN07016"  
/clone\_1lb="DrosBAC"  
/plasmid="pBelobAC11"  
/note="end : 17"

BASE COUNT 483 a 85 c 9 g 97 t 128 others  
ORIGIN

## Query Match

Best Local Similarity 35.8%; Pred. No. 16;  
Matches 261; Conservative 53; Mismatches 412; Indels 3; Gaps 1;

```

QY 12 AGACAACTAAAGTAATGAAAGAGATGCTGTAGAAAAGCAATATCATGAG 71
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 7 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 66

QY 72 TAATGTGTGACATTTGTGGGACCAAGAAATATAAGACATTTGATGCTGTATCC 131
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 67 AGAAMAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 126

QY 132 TCGATGAGCTCAAAAGTCTCTACACCCGATAGAAACCTTAAGCAATGTCAAGT 191
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 127 MAACGGGAACCNACCTACACCMCTCMNMAAAAAAAAAAATTAATAAAGCAACMTATA 186

QY 132 TTGCATTCCTCAGTGCATTAATGCAAAATTAAGATATCATGATGATAGCACTATGC 251
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 187 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAC 246

QY 252 ATCATATCATGCTCTCTCAACCTATTCATCTCTACTCATATAGATATCTTCAGCT 311
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 247 CACMCCTANNAAMAMCCCAAMATMAAAAAAAAAAAAAAAAAAGANNAAMAAAAA 306

QY 312 AAATGTTAGACATAACCCATTAAGTCACTGATGAGTATTAAGCGCTGACACATGAC 371
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 307 AAAAAAAAAAAAAAAAAAAAAAAAAACMAAMCAACNTANACCANNAAAAAAAAAA 366

QY 372 AATCAGACTCAGCAGAGATTAAGCAAAATGATGTGATCAATAAACCCAGGCTAAT 431
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 367 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANNA 426

QY 432 GTCAATTTGCAAAAAGAGAGAGCTTTAAGACAAGCATGATCACAATAA--TTTAC 488
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 427 AKNNNNNNCAANNNTAAAAAACAATAAATAAATAAATAAATAAATAAATAAATAA 486

QY 488 TTGCCCTTGTGCAAAAAGAGAGGCTTTAATTAATTCATGCTATTTGCAAAAAGAA 548
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 487 CTAATTTCAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 546

QY 549 GAGAGAAAGAACACATATGCTGCTGCAATTAATTAATGATGCTGCTCATATATTC 608
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 547 CAAAAAATMAACAAAAATCTCTTTTATTAATTAATTAATTAATTAATTAATTA 606

QY 609 ATCCACCTTGTGTACACACACTCATATATTAATTAAGAGTCACTTCAGCTGTGACATTA 668
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 607 ATTCATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 666

QY 669 ACAACCTATCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 728
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 667 AAAAAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 726

QY 729 TTCTCATTT 737

```

Db

727 YAAATWTV 735

## RESULT 15

B0595969

LOCUS

## DEFINITION

B0595969 563 bp mRNA linear EST 24-JUN-2002

PFEEST02b2603.Y1 Plasmodium falciparum 3D7 asexual cDNA Plasmodium

falciparum 3D7 cDNA 5', mRNA sequence.

## ACCESSION

B0595969

B0595969.1 GI:21542695

## VERSION

B0595969.1

## KEYWORDS

EST.

## SOURCE

Plasmodium falciparum 3D7

## ORGANISM

Plasmodium falciparum 3D7

## REFERENCE

1 (bases 1 to 563)

Eukaryota; Alveolata; Apicomplexa; Hemosporidia; Plasmodium.

Marta, K., Cole, R., Chakrabarti, D., Haywood, R., Clifton, S., Pape, D.,

Marta, M., Hillier, L., Martin, J., Wylie, T., Dante, M., Theising, B.,

Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Jentes, E., Ronko, I.,

Tagarelskii, R., Belaygorod, L., Franklin, C., Carr, L., Grow, A.,

Maguire, L., Ritchie, J., Madkins, J., Kennedy, S., Levinso, D.,

Waterston, R., Wilson, R. and Sibley, D.

Washu Plasmodium EST Project

Unpublished

Contact: L. David Sibley

Washu Plasmodium EST Project

Washington University School of Medicine

444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@wustl.edu

Library was constructed by Debopam Chakrabarti DNA sequencing by:

Washington University Genome Sequencing Center for information on

obtaining a clone please contact: L. David Sibley

(sibley@borchm.wustl.edu), Washington University

Seq primer: -40bp from Gibco

High quality sequence stop: 423.

## FEATURES

source

Location/Qualifiers

1..563

/organism="Plasmodium falciparum 3D7"

/mol\_type="mRNA"

/db\_xref="taxon:36329"

/lab\_host="DH10B (Genesig, Invitrogen, Inc.)"

/clone\_1lb="Plasmodium falciparum 3D7 asexual cDNA"

/note="Vector: pBluescript SK plus; Site 1: EcoRI; Site 2:

XhoI; Library was constructed by Debopam Chakrabarti.

Total RNA samples were isolated from mixed stage

saponin(0.1%) lysed P. falciparum 3D7 infected

erythrocytes by the acidic guanidinium-phenol chloroform

method. The poly A+ RNA was isolated by the polyAT-tract

mRNA isolation system (Promega, WI) using streptavidin

magnisphere particles. Directional cDNA libraries were

constructed by oligo d(T) priming of poly(A)+ RNA (5mg)

into EcoRI and XhoI sites of 1 ZapII vector using the Zap

cDNA synthesis kit (Stratagene, CA). The average size of the

cDNA inserts in the library was between 1.0 and 1.5kb.

Clones were mass excised using the exsist helper phage

(Stratagene), the phagelids were precipitated with PEG

8000 and extracted with phenol/chloroform. Phagelid DNA

was electroporated into DH10B cells."

BASE COUNT 347 a 35 c 71 g 110 t

## ORIGIN

## Query Match

Best Local Similarity 50.7%; Pred. No. 19;  
Matches 113; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

```

QY 359 GTGACATGACAAATTCACAGCTCAGCAAGTAAAGCAAAATGATGCTAATTAAC 418
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 211 GTCTTAAGGATTAATGAGAAATAATGAAAGATTAATAAAGATCAACAATAAT 270

QY 419 TCCAGAGCTATATGTCATATTCACAAAGAGAGAGCTTAATAAGCATGATCAGC 478
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```



Db 271 GAAAAATTAATTAATTAATGAATAAATACAGATCTGAAGAGAACAAAAATGAC 330  
QY 479 AAAATTCACCTTGCCTTCTGTCACAAAAGAGAGGCTTTACATTAATCATGTCATATT 538  
Db 331 AAAGAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 390  
QY 539 GCAAAAGAAAG 581  
Db 391 ATGAG 433

Search completed: September 14, 2003, 02:18:03  
Job time : 1029.62 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 13, 2003, 03:12:00 ; Search time 34.8026 Seconds  
(without alignments)  
9879.649 Million cell updates/sec

Title: US-09-847-232a-26

Perfect score: 779

Sequence: 1 ttctgtatgaagaaac.....ttcattagctcaaacac 779

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapect 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 113956

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued Patents\_NA.\*  
2: /cgn2\_6/prodata/2/ina/5A.COMB.seq:\*  
3: /cgn2\_6/prodata/2/ina/5B.COMB.seq:\*  
4: /cgn2\_6/prodata/2/ina/6A.COMB.seq:\*  
5: /cgn2\_6/prodata/2/ina/6B.COMB.seq:\*  
6: /cgn2\_6/prodata/2/ina/PCBUS.COMB.seq:\*  
7: /cgn2\_6/prodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	763.4	98.0	839 1	US-08-142-393-1
2	763.4	98.0	1240 1	US-08-142-393-2
3	39.6	5.1	7218 1	US-08-232-463-14
4	36.4	4.7	2965 2	US-08-460-570-1
5	36.4	4.7	2965 2	US-08-460-570-2
6	36.4	4.7	2965 3	US-08-286-870A-1
7	36.4	4.7	2965 3	US-08-286-870A-2
8	36.2	4.6	8585 1	US-08-030-096-3
9	36.2	4.6	58407 4	US-08-916-421B-2
10	35.2	4.5	3515 2	US-08-391-743A-1
11	35.2	4.5	3515 4	US-09-143-155-1
12	34.4	4.4	658 3	US-08-998-416-595
13	34	4.4	1678 1	US-08-261-677-10
14	34	4.4	1678 1	US-08-384-556A-6
15	34	4.4	1678 2	US-08-331-355A-10
16	34	4.4	1678 5	PCR-US94-12364-10
17	34	4.4	1678 5	PCR-US95-07753-6
18	34	4.4	2327 3	US-09-157-077-1
19	34	4.4	2796 1	US-08-261-677-8
20	34	4.4	2796 1	US-08-384-556A-4
21	34	4.4	2796 2	US-08-331-355A-8
22	34	4.4	2796 3	US-08-777-147-5
23	34	4.4	2796 3	US-09-157-077-5
24	34	4.4	2796 3	PCR-US94-12364-8
25	34	4.4	2796 5	PCR-US95-07753-4
26	34	4.4	148567 5	US-09-801-876B-3
27	34	4.4	152331 3	US-09-128-155-16

28	34	4.4	176373 3	US-09-128-155-17	Sequence 17, Appl
29	34	4.4	1230025 4	US-09-198-452A-1	Sequence 1, Appl
30	33.6	4.3	2494 4	US-09-016-434-1383	Sequence 1383, Ap
31	33.6	4.3	11091 4	US-09-134-001C-2243	Sequence 2243, Ap
32	33.6	4.3	1664976 4	US-08-916-421B-1	Sequence 1, Appl
33	33.4	4.3	294 4	US-09-328-352-2400	Sequence 2400, Ap
34	33.4	4.3	90541 4	US-09-759-359A-3	Sequence 3, Appl
35	33.2	4.3	879 4	US-09-134-001C-1305	Sequence 1305, Ap
36	33.2	4.3	19124 2	US-08-487-825B-13	Sequence 13, Appl
37	33.2	4.3	319608 4	US-09-539-335D-1	Sequence 1, Appl
38	33.2	4.3	319608 4	US-09-679-409-1	Sequence 1, Appl
39	33.2	4.3	1664976 4	US-08-916-421B-1	Sequence 1, Appl
40	32.8	4.2	3460 1	US-08-312-312A-1	Sequence 1, Appl
41	32.8	4.2	55827 4	US-09-813-133A-3	Sequence 3, Appl
42	32.8	4.2	202001 4	US-09-734-674-3	Sequence 3, Appl
43	32.6	4.2	546 4	US-08-328-352-992	Sequence 992, App
44	32.4	4.2	114 1	US-08-257-073-124	Sequence 124, App
45	32.4	4.2	114 1	US-08-257-073-125	Sequence 125, App

#### ALIGNMENTS

RESULT 1  
US-08-142-393-1  
; Sequence 1, Application US/08142393  
; Patent No. 5516668  
; GENERAL INFORMATION:  
; APPLICANT: Maruta, Yoshiyuki  
; TITLE OF INVENTION: Method for Decreasing Seed Storage  
; NUMBER OF INVENTIONS: 2  
; CORRESPONDENCE ADDRESSES:  
; ADDRESS: Birch, Stewart, Kolasch & Birch  
; STREET: 810 Gatehouse Road, Suite 500 East  
; CITY: Falls Church  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22042  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/142,393  
; FILING DATE: 24-NOV-1993  
; CLASSIFICATION: 800  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Murphy Jr., Gerald M.  
; REGISTRATION NUMBER: 28,977  
; REFERENCE/DOCKET NUMBER: 760-175P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)205-8000  
; TELEFAX: (703)205-8050  
; TELEX: 248345  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 839 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHEICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Oryza sativa  
; US-08-142-393-1  
Query Match 98.0%; Score 763.4; DB 1; Length 839;  
Best Local Similarity 99.6%; Pred. No. 2.9e-204;  
Matches 776; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

QY	1	TTCTGGATGACGACCAAACTCTAAAGTAAATGAAGAAGAGTGGGTCTTGAAAAGGAAC	60
Db	63	TTCTGGATGACGACCAAACTCTAAAGTAAATGAAGAAGAGTGGGTCTTGAAAAGGAAC	122
QY	61	AAATCATGAGTAATGTGTGACATTATAGGACCAAGAAATATAAAGACATTGTGATGA	120
Db	123	AAATCATGAGTAATGTGTGACATTATAGGACCAAGAAATATAAAGACATTGTGATGA	182
QY	121	GTGTGTATCTCGATAGAGCTTCAAAGTTCCTCACGCCGATAGAAAACCTTAGCA	180
Db	183	GTGTGTATCTCGATAGAGCTTCAAAGTTCCTCACGCCGATAGAAAACCTTAGCA	242
QY	181	ATGAGCAAGTTCATCTCCACATGACATATGCAAAATTAAGATATCATCATGACATA	240
Db	243	ATGAGCAAGTTCATCTCCACATGACATATGCAAAATTAAGATATCATCATGACATA	302
QY	241	GCAACATCATGATATATCATAGCCCTCTCCACACTTATCTCATCTCATCATATAG	300
Db	303	GCAACATCATGATATATCATAGCCCTCTCCACACTTATCTCATCTCATCATATAG	362
QY	301	TATCTTCACATCAATATGTTTAAACATAAACCATAACATCAAGTTGATGAGATATAGGGGT	360
Db	363	TATCTTCACATCAATATGTTTAAACATAAACCATAACATCAAGTTGATGAGATATAGGGGT	422
QY	361	GACACATGACAAATTCACAGACTCAAGCAAGATPAAGCAAAATGATGTGATCAATAAACTC	420
Db	423	GACACATGACAAATTCACAGACTCAAGCAAGATPAAGCAAAATGATGTGATCAATAAACTC	482
QY	421	CAGAGCTAATATGCTATTTGCCAAAAGAGAGAGCTTATATAGACAAAGCATGATCAACAA	480
Db	483	CAGAGCTAATATGCTATTTGCCAAAAGAGAGCTTATATAGACAAAGCATGATCAACAA	542
QY	481	AAATTCACATGCTTCCTTGTTGTTCCAAAAGAGAGAGCTTATATAGATATGATCATATTC	540
Db	543	AAATTCACATGCTTCCTTGTTGTTCCAAAAGAGAGAGCTTATATAGATATGATCATATTC	602
QY	541	AAAAGAAAAGAGAAAGAACACACACATAGCTGGCTCATATATATATCTGTATGTCCAT	600
Db	603	AAAAGAAAAGAGAAAGAACACACACATAGCTGGCTCATATATATATCTGTATGTCCAT	662
QY	601	CATATTCATCCACCTTGGTGTACACACTCATATATCATATAGAGTCCACTCCAGTCT	660
Db	663	CATATTCATCCACCTTGGTGTACACACTCATATATCATATAGAGTCCACTCCAGTCT	722
QY	661	GGACATTAACAACACTCATCTTAACATTTAGATGACAAAGACCTTATCTCATATPAATG	720
Db	721	GGACATTAACAACACTCATCTTAACATTTAGATGACAAAGACCTTATCTCATATPAATG	780
QY	721	CACAGATGATTTCTCATTTGTTTCCAAAAGATTCACATTCATATAGTCCCTAATCAAC	778
Db	781	CACAGATGATTTCTCATTTGTTTCCAAAAGATTCACATTCATATAGTCCCTAATCAAC	838

RESULT 2

US-08-142-393-2

Sequence 2, Application US/08142393

Patent No. 5516668

GENERAL INFORMATION:

APPLICANT: Maruta, Yoshiyuki

TITLE OF INVENTION: Method for Decreasing Seed Storage

TITLE OF INVENTION: Proteins and for Transforming Plants

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: Birch, Stewart, Kolasch & Birch

STREET: 8110 Gatehouse Road, Suite 500 East

CITY: Falls Church

STATE: Virginia

COUNTRY: U.S.A.

ZIP: 22042

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

1	SOFTWARE: PatentIn Release #1.0, Version #1.25
2	CURRENT APPLICATION DATA:
3	APPLICATION NUMBER: US/06/142,393
4	FILING DATE: 24-NOV-1993
5	CLASSIFICATION: 800
6	ATTORNEY/AGENT INFORMATION:
7	NAME: Murphy Jr., Gerald M.
8	REGISTRATION NUMBER: 28,977
9	REFERENCE/DOCKET NUMBER: 760-175P
10	TELECOMMUNICATION INFORMATION:
11	TELEPHONE: (703)205-8000
12	TELEFAX: (703)205-8050
13	TELEX: 248345
14	INFORMATION FOR SEQ ID NO: 2:
15	SEQUENCE CHARACTERISTICS:
16	LENGTH: 1240 base pairs
17	TYPE: nucleic acid
18	STRANDEDNESS: double
19	TOPOLOGY: linear
20	MOLECULE TYPE: DNA (genomic)
21	HYPOHETICAL: NO
22	ANTI-SENSE: NO
23	ORIGINAL SOURCE:
24	ORGANISM: Oryza sativa
25	US-06-142-393-2
26	
27	Query Match
28	Best Local Similarity 98.0%; Score 763.4; DB 1; Length 1240;
29	Matches 776; Conservative 0; Mismatches 1; Indels 2; Gaps 1.
30	
31	1 TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 60
32	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 120
33	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 180
34	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 240
35	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 300
36	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 360
37	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 420
38	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 480
39	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 540
40	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 600
41	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 660
42	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 720
43	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 780
44	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 840
45	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 900
46	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 960
47	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 1020
48	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 1080
49	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 1140
50	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 1200
51	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 1260
52	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 1320
53	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 1380
54	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 1440
55	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 1500
56	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 1560
57	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 1620
58	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 1680
59	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 1740
60	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 1800
61	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 1860
62	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 1920
63	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 1980
64	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 2040
65	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 2100
66	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 2160
67	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 2220
68	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 2280
69	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 2340
70	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 2400
71	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 2460
72	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 2520
73	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 2580
74	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 2640
75	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 2700
76	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 2760
77	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 2820
78	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 2880
79	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 2940
80	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 3000
81	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 3060
82	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 3120
83	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 3180
84	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 3240
85	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 3300
86	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 3360
87	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 3420
88	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 3480
89	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 3540
90	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 3600
91	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 3660
92	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 3720
93	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 3780
94	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 3840
95	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 3900
96	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 3960
97	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 4020
98	TTCTGTACTGACAGCAAAACCTAAAGTAAAGAGAGATGTGGTTAGAAAAGAAAC 4

QY	661	GGACATTAAACAAACATCTCTATACATTAAAGCAAGCCCTTATCTACGTAATAAGT	720
Db	721	GGACATTAAACAAACCTATCTTACATTTAAATCGAAGCCCTTATCTACGTAATAAGT	780
QY	721	CAGAGATTTCTCATTTGTTCTTCACAAAAGCATTCAGTTGTAATGCTCTCAACAAAC	779
Db	781	CAGAGATTTCTCATTTGTTCTTCACAAAAGCATTCAGTTGTAATGCTCTCAACAAAC	839

### RESULT 3

```

US-08-232463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSER: Foley & Lathier
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZ9pt-F15
US-08-232463-14

```

Query Match 5.1%; Score 39.6; DB 1; Length 7218;  
Best Local Similarity 5.0%; Pred. No. 0.18;  
Matches 15; Conservative 164; Mismatches 123; Indels 0; Gaps 0;

```

OY      312  AAAAGTTAGAACATAAACCATAAGTCACGTTGAGTAATATAGGCGTGACACATGAA 371
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1341  RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 128
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY      372  AATCACGAGTCACGACGAGATTAAGCAAAATGATGTACATAAATCCAGAGCATAT 431
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1281  RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1222
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY      432  GTCATATTGCAAAAAGAGGAGAGCTTTAAAGACAGGCGATGCACAAAAATTCCTGG 491
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```
Db      1221 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1102
QY      492 COTTGTCGCAAAAAGAGGGCTTTACATTATCATGTGCATATTGCAAAAGAAGAG 551
        : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1161 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1102
QY      552 AGAAGAACAACACTGTCGTCATTTTACATTCCTGTAATGCCATCATTATTCAC 611
        : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1101 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1042
QY      612 CA 613
Db      1041 CA 1040
```

## RESULT 4

US-08-460-570-1  
Sequence 1, Application US/08460570  
Patent No. 5965797  
GENERAL INFORMATION:  
APPLICANT: BLENK, ROBERT G.  
APPLICANT: ELY, SUSAN  
APPLICANT: TAILOR, RAVINDRA H.  
APPLICANT: TIPEET, JANET M.  
TITLE OF INVENTION: BACTERIAL GENES  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CUSHMAN DABNEY & CUSHMAN, L.L.P.  
STREET: 1100 New York Avenue, N.W.  
CITY: Washington  
STATE: D. C.  
COUNTRY: U.S.A.  
ZIP: 20005-3918  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/460,570  
FILING DATE: 02-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/520,228  
FILING DATE: 09-MAY-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 8910624.9  
FILING DATE: 09-MAY-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: KOKULIS, Paul N.  
REGISTRATION NUMBER: 16,773  
REFERENCE/DOCKET NUMBER: 218975/PS.35271/US/N  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 861-3000  
TELEFAX: (202) 822-0944  
TELEX: 6714627 CUSH  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2965 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: both  
TOPOLOGY: linear  
US-08-460-570-1

		4.78;	Score	36.4;	DB	2;	Length	2965;	
Query Match Similarity		50.66;	Pred.	No.	1;				
Best Local									
Matches	88;	Conservative	0;	Mismatches	86;	Indels	0;	Gaps	0
OY	394	AAGCAAAATGATGTGTGTCATAAATCTCGACGACTATTTGCATATTGCCAAAAGAGAGA	453						
Db	288	AAAGTATATTTTGATGTAATAAATATTATTTGTGAATTTTAAATAATAGTATTAAGGAGG	347						
OY	454	GCTTATTAAGACAACGACATGACTCACAAAAATTCATTGGCTTTCGTGTCAAAGAGAGAG	513						



US-08-286-870A-2  
; Sequence 2, Application US/08286870A  
; Patent No. 6063605  
; GENERAL INFORMATION:  
; APPLICANT: ELY, S  
; APPLICANT: TAILOR, RH  
; APPLICANT: TIPPETT, JM  
; APPLICANT: BLENN, RG  
; TITLE OF INVENTION: BACTERIAL GENES  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CUSHMAN DABRY & CUSHMAN  
; ADDRESSEE: Intellectual Property Group of  
; ADDRESSEE: PILLSBURY, MADISON & SUTRO LLP  
; STREET: 1100 New York Avenue, N.W.  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005-3918  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/286,870A  
; FILING DATE: 05-AUG-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/520228  
; FILING DATE: 09-MAY-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 8910624.9  
; FILING DATE: 09-MAY-1989  
; ATTORNEY/AGENT INFORMATION:  
; NAME: PAUL N. KOKULIS  
; REGISTRATION NUMBER: 16,773  
; REFERENCE/DOCKET NUMBER: 70608/220720  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 861-3000  
; TELEFAX: (202) 822-0944  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2965 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: both  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..2965  
US-08-286-870A-2  
Query Match 4.7%; Score 36.4; DB 3; Length 2965;  
Best Local Similarity 50.6%; Pred. No. 1;  
Matches 88; Conservative 0; Mismatches 86; Indels 0; Gaps 0;  
QY 394 AACCAATGATGTGTCATPAACATCCAGACTATATGTCATATTCGCAAAAAGAGAGA 453  
DB 288 AAGTATATTTTGTATGATATAAATAATATATCTGAATAATTAAATGATAGTGAGAG 347  
QY 454 GCTTAAAGACAGGACATGACTCACAATAATTCACCTTCCTTGTGTCTAAAAAGAGAG 513  
DB 348 GATTATATAGAAATCAAGATCAAGATTAAGCATCAAGTTTTCATAGCAATGCGAAGT 407  
QY 514 GCTTACATATTCATGTCATATTCGCAAAAAGAGAGAAAGACACACACA 567  
DB 408 AGATTAATCTCTAGGATTCCTACTAAAAAATGAACACATATTAATTAATAA 461

RESULT 8  
US-08-030-096-3/C  
; Sequence 3, Application US/08030096

; Patent No. 5426041  
; GENERAL INFORMATION:  
; APPLICANT: Fabijanski, Steven F.  
; APPLICANT: Arnison, Paul G.  
; TITLE OF INVENTION: BINARY CRYPTOCYTOXIC METHOD OF HYBRID  
; SEED PRODUCTION  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington, D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/030,096  
; FILING DATE: 22-MAR-1993  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/556,917  
; FILING DATE: 20-JUL-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCY/CA91/00255  
; FILING DATE: 22-JUL-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 33229/164/PIHI  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 8585 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: join(308..370, 1136..1261, 6369..6428, 7198..7353)  
US-08-030-096-3  
Query Match 4.6%; Score 36.2; DB 1; Length 8585;  
Best Local Similarity 48.3%; Pred. No. 1.7;  
Matches 101; Conservative 0; Mismatches 108; Indels 0; Gaps 0;  
QY 535 TATTCGCAAG 594  
DB 8410 TATTAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 8351  
QY 595 GTCCATCATATTCATCCACCTTGTGTCACACCTTCATATATCAATAGAGTCACTC 654  
DB 8350 GTAAAGAGCACTACATGAGAGTGTATATATCTATTAATTCGGTATTTTGTGTGTT 8291  
QY 655 AGCTGAGACATTAACAACTCATCTATTAAGTTAGTAGAGAGAGAGAGAGAGAGAG 714  
DB 8290 TAGBACTCTCAAGGTGTGAAGCTTTAAGATTTAATGAACATGTTTACTTAACT 8231  
QY 715 TAAATGACAGATGATTCCTATTTCTTCT 743  
DB 8230 TCTTTGCCACATATTTCTCTCTTCTTACT 8202

RESULT 9  
US-08-916-421B-2/C  
; Sequence 2, Application US/08916421B  
; Patent No. 6503729

TELEPHONE: (617) 227-7400  
TELEFAX: (617) 742-4214

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match 4.5%; Score 35.2; DB 4; Length 3515;





QY 503 AAAAGAGAGGCGCTTACATTATCCATGTCATATTGCAAAAGAGAGAGACCA 562  
|||||  
Db 718 TAAAGAGATGATGATTAATCTACTAATAATATGAGAGTCAAAATACAGAGAAACACT 777

QY 563 CA 564  
||  
Db 778 CA 779

RESULT 14  
US-08-384-556A-6  
; Sequence 6, Application US/08384556A  
; Patent No. 5750870  
; GENERAL INFORMATION:  
; APPLICANT: Mathews, Helena V  
; APPLICANT: Bestwick, Richard K  
; APPLICANT: Petro, Adolph J  
; TITLE OF INVENTION: Plant Genetic Transformation  
; TITLE OF INVENTION: Methods and Transgenic Plants  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dehlinger & Associates  
; STREET: 350 Cambridge Avenue, Suite 250  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/384,556A  
; FILING DATE:  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/263,900  
; FILING DATE: 17-JUN-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fabian, Gary R.  
; REGISTRATION NUMBER: 33,875  
; REFERENCE/DOCKET NUMBER: 4257-0010.30  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 324-0880  
; TELEFAX: (415) 324-0960  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1678 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; INDIVIDUAL ISOLATE: E4 tomato promoter /  
; INDIVIDUAL ISOLATE: Adometase gene DNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1174..1629  
; US-08-384-556A-6

Query Match 4.4%; Score 34; DB 1; Length 1678;  
Best Local Similarity 54.9%; Pred. No. 3.8;  
Matches 67; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 443 AAAAGAGAGGCTTATAGACAGGACATGCTCCACAAAATTCCTTGGCTTGGTGC 502  
|||||  
Db 656 AAACGAGATATCCAAAATTTCAAAATTTAATGATGATTTGGTCATGATCTACGTCGA 717  
QY 503 AAAAGAGAGGCGCTTACATTATCCATGTCATATTGCAAAAGAGAGAGACCA 562

Db 718 TAAAGAGATGATGATTAATCTACTAATAATATGAGAGTCAAAATACAGAGAAACACT 777  
|||||  
QY 563 CA 564  
||  
Db 778 CA 779

RESULT 15  
US-08-331-355A-10  
; Sequence 10, Application US/08331355A  
; Patent No. 5859330  
; GENERAL INFORMATION:  
; APPLICANT: Bestwick, Richard K  
; APPLICANT: Petro, Adolph J  
; TITLE OF INVENTION: Regulated Expression of Heterologous  
; TITLE OF INVENTION: Genes in Plants and Transgenic Fruit  
; TITLE OF INVENTION: with a Modified Ripening Phenotype  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dehlinger & Associates  
; STREET: 350 Cambridge Avenue, Suite 250  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/331,355A  
; FILING DATE: 27-OCT-1994  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/261,677  
; FILING DATE: 17-JUN-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/046,583  
; FILING DATE: 09-APR-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/255,833  
; FILING DATE: 08-JUN-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/613,858  
; FILING DATE: 12-DEC-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 448,095  
; FILING DATE: 12-DEC-1989  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Evans, Susan T.  
; REGISTRATION NUMBER: 38,443  
; REFERENCE/DOCKET NUMBER: 4257-0011.30  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 324-0880  
; TELEFAX: (415) 324-0960  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1678 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; INDIVIDUAL ISOLATE: Figure 25 - E4 tomato promoter /  
; INDIVIDUAL ISOLATE: Adometase gene DNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1174..1629  
; US-08-331-355A-10





```
|||||
Db 358 GCAAGTTGGCATTCCTCCACTGACATTAATGCAAAATAGATATCATGATGACATACGAA 417
Qy 245 CTCATGCAATATATCATGCTCTCTCTCAACCTATTCATCTCTACTCTACTATGATATC 364
Db 418 CTCATGCAATATATCATGCTCTCTCTCAACCTATTCATCTCTACTCTACTATGATATC 477
Qy 305 TTCAGCTAAATGTTAGACATTAACCCATTAAGTACAGCTTTGATGATATTAAGGGGTACA 364
Db 478 TTCAGCTAAATGTTAGACATTAACCCATTAAGTACAGCTTTGATGATATTAAGGGGTACA 537
Qy 365 CATGCAAAATPCACAGACTCAAGCAGATTAAGCAAAATGATGTACATTAATACCTCAGA 424
Db 538 CATGCAAAATPCACAGACTCAAGCAGATTAAGCAAAATGATGTACATTAATACCTCAGA 597
Qy 425 GCTATATGTCATATTTGCAAAAAGAGAGGCTTTAAGACAAAGGCTGACACAAAAT 484
Db 598 GCTATATGTCATATTTGCAAAAAGAGAGGCTTTAAGACAAAGGCTGACACAAAAT 657
Qy 485 TCACCTTGCTTCGTGTCAAAAAGAGAGGCTTTAAGATTAATCATGATGATATGCAAAA 544
Db 658 TCACCTTGCTTCGTGTCAAAAAGAGAGGCTTTAAGATTAATCATGATGATATGCAAAA 717
Qy 545 GAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 604
Db 718 GAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 777
Qy 605 ATTCATCACCTTCCTGCTGCTCAACACTTCATTAATCATTAAGATCACTTCAGCTGTGAC 664
Db 778 ATTCATCACCTTCCTGCTGCTCAACACTTCATTAATCATTAAGATCACTTCAGCTGTGAC 837
Qy 665 ATTAACAACCTTCATTAATCATTAAGATCACTTCAGCTGTGAC 724
Db 838 ATTAACAACCTTCATTAATCATTAAGATCACTTCAGCTGTGAC 897
Qy 725 ATGATTTTCATTTGTTTCTCAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 779
Db 898 ATGATTTTCATTTGTTTCTCAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 952
```

## RESULT 2

```
US-10-076-816-15
; Sequence 15, Application US/10076816
; Publication No. US20030056244A1
; GENERAL INFORMATION:
; APPLICANT: Huang, Ning
; APPLICANT: Rodriguez, Raymond
; TITLE OF INVENTION: Feed Additive Compositions and Methods
; FILE REFERENCE: 50665-8021.US00
; CURRENT APPLICATION NUMBER: US/10/076,816
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/269,188
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/847,232
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/266,929
; PRIOR FILING DATE: 2001-02-06
; PRIOR APPLICATION NUMBER: US 60/201,182
; PRIOR FILING DATE: 2000-05-02
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 786
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Rice Gtl promoter and Gtl leader coding sequence
US-10-076-816-15
```

```
Query Match 90.6%; Score 706; DB 14; Length 786;
Best Local Similarity 99.3%; Pred. No. 1,5e-172;
Matches 709; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
Qy 66 CATGAGTAATGTGAGGATTAATGGACACGAAATATAAGACATTTGATGAGTGT 125
Db 1 CATGAGTAATGTGAGGATTAATGGACACGAAATATAAGACATTTGATGAGTGT 60
Qy 126 GTATCCGATGAGGCTCAAAAGTTCTCTACCCCGGATTAAGAAACCTTAAGCAATG 185
Db 61 GTATCCGATGAGGCTCAAAAGTTCTCTACCCCGGATTAAGAAACCTTAAGCAATG 120
Qy 186 CAAAGTTGATTTCCACTGACATTAATGCAAAATTAAGTATCATGATGATGACATG 245
Db 121 CAAAGTTGATTTCCACTGACATTAATGCAAAATTAAGTATCATGATGATGACATG 180
Qy 246 TCATGATATATATCATGCTCTCTCAACCTATTCATTCCTACTCATATATATATCT 305
Db 181 TCATGATATATATCATGCTCTCTCAACCTATTCATTCCTACTCATATATATATCT 240
Qy 306 TCAGCTTAATGTTAGACATTAACCCATTAAGTACGTTGATGATGATTTAGGGGTGAC 365
Db 241 TCAGCTTAATGTTAGACATTAACCCATTAAGTACGTTGATGATGATTTAGGGGTGAC 300
Qy 366 ATGACAATGACAGACTACAGACAGATTAAGCAAAATGATGTACATTAATACCTCAG 425
Db 301 ATGACAATGACAGACTACAGACAGATTAAGCAAAATGATGTACATTAATACCTCAG 360
Qy 426 CTATATGTCATATTTGCAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 485
Db 361 CTATATGTCATATTTGCAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Qy 486 CACTTGCTTCGTGTCAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 545
Db 421 CACTTGCTTCGTGTCAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
Qy 546 AAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 605
Db 481 AAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Qy 606 TTCATCACCTTCCTGCTGCTCAACACTTCATTAATCATTAAGATCACTTCAGCTGTGAC 665
Db 541 TTCATCACCTTCCTGCTGCTCAACACTTCATTAATCATTAAGATCACTTCAGCTGTGAC 600
Qy 666 TTAACAACCTTCATTAATCATTAAGATCACTTCAGCTGTGAC 725
Db 601 TTAACAACCTTCATTAATCATTAAGATCACTTCAGCTGTGAC 660
Qy 726 TGATTTCTCATTTGTTTCTCAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 779
Db 661 TGATTTCTCATTTGTTTCTCAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 714
```

## RESULT 3

```
US-10-077-381-15
; Sequence 15, Application US/10077381
; Publication No. US20030074700A1
; GENERAL INFORMATION:
; APPLICANT: Huang, Ning
; APPLICANT: Rodriguez, Raymond
; TITLE OF INVENTION: Expression of Human Milk Proteins in
; FILE REFERENCE: 50665-8022.US00
; CURRENT APPLICATION NUMBER: US/10/077,381
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/269,199
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/847,232
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/266,929
; PRIOR FILING DATE: 2001-02-06
; PRIOR APPLICATION NUMBER: US 60/201,182
; PRIOR FILING DATE: 2000-05-02
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
```

```

: SEQ ID NO 15
:
: LENGTH: 786
:
: TYPE: DNA
:
: ORGANISM: Artificial Sequence
:
: FEATURE:
:
: OTHER INFORMATION: Rice Gtl promoter and Gtl leader coding sequence
US-10-077-381-15

```

Query Match	90.6%;	Score 706;	DB 14;	Length 786;
Best Local Similarity	99.3%;	Pred. No. 1.5e-172;		
Matches 709; conservative	0;	Mismatches 5;	Indels 0;	Gaps 0;

QY	66	CATGAGTAAAGTGTGAGCATTTTGGGACCCGAAATTAATAATTTGAGAGTCT	125
Db	1	CATGAGTAAAGTGTGAGCATTTTGGGACCCGAAATTAATAATTTGAGAGTCT	60
QY	126	GTATCTCGATGAGGCTCAAAAGTTCTCTCAGCCCGGAAATGAACCTTACGAAATG	185
Db	61	GTATCTCGATGAGGCTCAAAAGTTCTCTCAGCCCGGAAATGAACCTTACGAAATG	120
QY	186	CAAGTTTGGATTTCTCCATGTCACATTAATGCAAAATAGATATACGATGACATGAC	245
Db	121	CAAGTTTGGATTTCTCCATGTCACATTAATGCAAAATAGATATACGATGACATGAC	180
QY	246	TCAATGATCATATCATGCTCTCTCAACCTATTCAATTCCTACTCATCTAATAGATCT	305
Db	181	TCAATGATCATATCATGCTCTCTCAACCTATTCAATTCCTACTCATCTAATAGATCT	240
QY	306	TCAGCTAAATGTTAGAACATTAACCCATATGATCAGCTTTGAGATATATAGGGTACAC	365
Db	241	TCAGCTAAATGTTAGAACATTAACCCATATGATCAGCTTTGAGATATATAGGGTACAC	300
QY	366	ATGACAATATCAGACTCAGCAGATTAAGCAAAATGATGTACATTAATCAACCCAGAG	425
Db	301	ATGACAATATCAGACTCAGCAGATTAAGCAAAATGATGTACATTAATCAACCCAGAG	360
QY	426	CTAATATGTCATATGCAAAAAGAGAGGCTTTATAGACAAGGCTATGCTACAAAAATT	485
Db	361	CTAATATGTCATATGCAAAAAGAGAGGCTTTATAGACAAGGCTATGCTACAAAAATT	420
QY	486	CACATGGCTTTCGTGTCAAAAAGAGAGGCTTTACATATTCAGATATATTTGCAAAAG	545
Db	421	CACATGGCTTTCGTGTCAAAAAGAGAGGCTTTACATATTCAGATATATTTGCAAAAG	480
QY	546	AAAGAGAAAGAAACACACCAATGCTGCTGCAATTTATCATATCTGTATGTCAATATTA	605
Db	481	AAAGAGAAAGAAACACACCAATGCTGCTGCAATTTATCATATCTGTATGTCAATATTA	540
QY	606	TTCAATCCACCTTTCGTGTACACACATTCATATATCATTAAGAGTCACTCAAGTGTGACA	665
Db	541	TTCAATCCACCTTTCGTGTACACACATTCATATATCATTAAGAGTCACTCAAGTGTGACA	600
QY	666	TTTAACAACCTATCTTAAACATTTAGATGACAAGAGCTTTATCTCACTATTAATGACAGA	725
Db	601	TTTAACAACCTATCTTAAACATTTAGATGACAAGAGCTTTATCTCACTATTAATGACAGA	660
QY	726	TCATTTCTCATTTGTTCTCCAAAAAGCATTCAGCTTCATTAATCTCTAATAAATAC	779
Db	661	TCATTTCTCATTTGTTCTCCAAAAAGCGCTGCTATTAATCTCTAATAAATAC	714

RESULT 4  
US-09-944-160-15  
Sequence 15, Application US/09944160  
Patent No. US20020174452A1  
GENERAL INFORMATION:  
APPLICANT: Lewis, No. US20020174452A1man  
APPLICANT: Davin, Laurence  
APPLICANT: .. Huang, Ming  
TITLE OF INVENTION: Monocot Seeds with Increased Lignin  
TYPE OF INVENTION: Content  
FILE REFERENCE: W09/117983  
CURRENT APPLICATION NUMBER: US/09/944,160

```

1 CURRENT FILING DATE: 2001-08-30
2 PRIOR APPLICATION NUMBER: US 60/230,632
3 PRIOR FILING DATE: 2000-09-07
4 NUMBER OF SEQ ID NOS: 49
5 SOFTWARE: FastSeq for Windows Version 4.0
6 SEQ ID NO 15
7 LENGTH: 4227
8 TYPE: DNA
9 ORGANISM: Artificial sequence
10 FEATURE:
11 OTHER INFORMATION: plasmid pBP1244 encoding divergent protein
12 US-09-944,160-15

```

```
Query Match      88.4%; Score 689; DB 10; Length 4227;
Best Local Similarity 100.0%; Pred. No. 8.6e-168;
Matches 689; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY	6	TCATGAGATATGCTGAGCATATGGGACACCAATAATATAAAGAACATTTTGATGGAGTGG	124
Db	238	TCATGAGATATGCTGAGCATATGGGACACCAATAATATAAAGAACATTTTGATGGAGTGG	297
QY	125	TGTAATCTCGCATAGGCTTCAAAAGTCTCTCCACCCGGGTAAAGAACCCCTTAAGCATCT	184
Db	298	TGTAATCTCGCATAGGCTTCAAAAGTCTCTCCACCCGGGTAAAGAACCCCTTAAGCATATP	357
QY	185	GCAAGATTTCATCTCCCACTGACATATGCAAAATATATCATCATGATGATAGCA	244
Db	358	GCAAGATTTCATCTCCCACTGACATATGCAAAATATATCATCATGATGATAGCA	417
QY	245	CTCATGATATATCATGCTCTCTCAACCTATTCATTTCTTACTCATATCAATATATC	304
Db	418	CTCATGATATATCATGCTCTCTCAACCTATTCATTTCTTACTCATATCAATATATC	477
QY	305	TTCAAGCTAAATGTTAAACATATTAACCCATATAGTCACAGTTGATGATTAAGGGGTGACA	364
Db	478	TTCAAGCTAAATGTTAAACATATTAACCCATATAGTCACAGTTGATGATTAAGGGGTGACA	537
QY	365	CATGACAAATTCACAGCTCAGCAAGATTAAGCAAAATATGATGTGTACATATTAACCTCAGA	424
Db	538	CATGACAAATTCACAGCTCAGCAAGATTAAGCAAAATATGATGTGTACATATTAACCTCAGA	597
QY	425	GCTATATGTATATTCGCAAAAAGAGAGAGCTTATATAGCAAGAGATGACATCCACAATAT	484
Db	598	GCTATATGTATATTCGCAAAAAGAGAGAGCTTATATAGCAAGAGATGACATCCACAATAT	657
QY	485	TCACCTGGCTTTGGTGTCAAAAAGAGAGAGGCTTTACATTAATATCATATTCATATGCAAAA	544
Db	658	TCACCTGGCTTTGGTGTCAAAAAGAGAGAGGCTTTACATTAATATCATATTCATATGCAAAA	717
QY	545	GAAAGAGAGAAAGAAACAAACACATATGCTGCTCAATTAATCATATCTGTATATGCTCATAT	604
Db	718	GAAAGAGAGAAAGAAACAAACACATATGCTGCTCAATTAATCATATCTGTATATGCTCATAT	777
QY	605	ATTCACTACACCTTCCGTGACACACCTCATATATCATATAGAGTCACTTCACGCTGGAC	664
Db	778	ATTCACTACACCTTCCGTGACACACCTCATATATCATATAGAGTCACTTCACGCTGGAC	837
QY	665	ATTAAACAACCTATCTTATACATTTAGATGCAAGAGGCTTATCTACATATTAATGCAAG	724
Db	838	ATTAAACAACCTATCTTATACATTTAGATGCAAGAGGCTTATCTACATATTAATGCAAG	897
QY	725	ATGATTTTCTCATTTGTTTCCACAAAAAC	753
Db	898	ATGATTTTCTCATTTGTTTCCACAAAAAC	926

RESULT5  
US-09-944-160-19  
; Sequence 19, Application US/099444160  
; Patent No. US20020174452A1  
; GENERAL INFORMATION:  
; APPLICANT: Lewis, No. US20020174452A1man  
; APPLICANT: Davin, Laurence

Query Match	88.4%	Score 689;	DB 10;	Length 5355;
Best Local Similarity	99.3%;	Pred. No. 9.6e-168;		
Matches 692; Conservative	0;	Mismatches 5;	Indels 0;	Gaps 0

Query Match	88.4%	Score 689;	DB 10;	Length 5355;
Best Local Similarity	99.3%;	Pred. No. 9,6e-168;		
Matches 662;	Conservative 0;	Mismatches 5;	Indels 0;	Gaps 0;
QY	65	TCATGAGTAAATGTGTGAGCATTTATGGACACACAAAATAAAAAGAACATTTGATGATCG	124	
Db	238	TCATGAGTAAATGTGTGAGCATTTATGGACACACAAAATAAAAAGAACATTTGATGATCG	297	
QY	125	TGTTATCTCCGATGAGGCTTCAAAAGTCTCCACCCGGATGAGAAACCTTAAGCAATCG	184	
Db	298	TGTTATCTCCGATGAGGCTTCAAAAGTCTCCACCCGGATGAGAAACCTTAAGCAATCG	357	
QY	185	GCAAGATTGCAATTCACACTGACATATGCAAAATAGATATCATCGATGACATAGCAA	244	
Db	358	GCAAGATTGCAATTCACACTGACATATGCAAAATAGATATCATCGATGACATAGCAA	417	
QY	245	CTCATGCAATATATCATGCTCTCTCAACCATTTCAATCTCTCATATCTACATATGATTC	304	
Db	418	CTCATGCAATATATCATGCTCTCTCAACCATTTCAATCTCTCATATCTACATATGATTC	477	
QY	305	TTTCAGCAATTAATGTTAAACAATATAACCAATAGACAGTTTGATGAGTTTATGGCGTGACA	364	
Db	478	TTTCAGCAATTAATGTTAAACAATATAACCAATAGACAGTTTGATGAGTTTATGGCGTGACA	537	
QY	365	CATGCAAAATACAGACTACAGACAGATATAAGCAAAATGATGTGACATATAAACTCCAGA	424	
Db	538	CATGCAAAATACAGACTACAGACAGATATAAGCAAAATGATGTGACATATAAACTCCAGA	597	
QY	425	GCTATATGCTATTTGCCAAAAGAGAGAGCTTATTAAGCAAGGATGACACCAAAAAT	484	
Db	598	GCTATATGCTATTTGCCAAAAGAGAGAGCTTATTAAGCAAGGATGACACCAAAAAT	657	
QY	485	TCACCTGGCTTTGCTGTCAAAAAGAGAGGCTTTACATTTATCOATGTCATATGCAAAA	544	
Db	658	TCACCTGGCTTTGCTGTCAAAAAGAGAGGCTTTACATTTATCOATGTCATATGCAAAA	717	
QY	545	GAAAGAGAGAAAGAACACACACATGCTGCTCAATTAATACATTTGTATGTCCATCAT	604	
Db	718	GAAAGAGAGAAAGAACACACATGCTGCTCAATTAATACATTTGTATGTCCATCAT	777	
QY	605	ATTTCATCCACCTTTGCGTACACACATTCATATATCTCTTAAGATGTCACACGTCGGA	664	
Db	778	ATTTCATCCACCTTTGCGTACACACATTCATATATCTCTTAAGATGTCACACGTCGGA	837	
QY	665	ATTTAACAACCTCTATCTATAACATTAGATGCAAGAGGCTTTATCTCACTATATAATGCAG	724	
Db	838	ATTTAACAACCTCTATCTATAACATTAGATGCAAGAGGCTTTATCTCACTATATAATGCAG	897	
QY	725	ATGATTTCTCATTTCTTCTTCACAAAAGACATTCAGTT	761	

Db 898 ATGATTTCTCATTTGTTTCTCACAAAAAGCGCGGCTT 934

RESULT 7  
US-09-944-160-7  
; Sequence 7, Application US/09944160  
; Patent No. US20020174452A1

1 APPLICANT: Lewis, No. US20020174452Alman  
 2 APPLICANT: Davine, Laurence  
 3 APPLICANT: .. Hang, Ning  
 4 TITLE OF INVENTION: Monocot Seeds with Increased Lignan  
 5 TITLE OF INVENTION: Content  
 6

Query Match 88.3%; Score 688; DB 10; Length 690;

QY	66	AATGAGTAAAGTGTGAGCTTATATGAGGACACAGAAATTAAGAAATTTGATGAGAGCCG	125
Dp	1	CATGAGTAAATGTGTGACCTTTATGGACACAGAAATTAAGAAATTTGATGAGAGCTG	60
QY	126	GTATCTCGATGAGGCTCAAAAGTTCTCTACCCCGAGTAAGAAACCTTAAAGCAATGTG	185
Dp	61	GTATCTCGATGAGGCTCAAAAGTTCTCTACCCCGAGTAAGAAACCTTAAAGCAATGTG	120
QY	186	CAAAAGTTGCATTTCTCCACTGACATATGCAAAATTAAGTATATCATGATGACATAGCAAC	245
Dp	121	CAAAAGTTGCATTTCTCCACTGACATATGCAAAATTAAGTATATCATGATGACATAGCAAC	180
QY	246	TCAAGCATCATATCAAGGCTCTCTCAACACATTCATTCCTACTCATGACATAGATATCG	305
Dp	131	TCAAGCATCATATCAAGGCTCTCTCAACACATTCATTCCTACTCATGACATAGATATCG	240
QY	306	TCAAGCTTAATGTTAGAACATTAACCCATTAAGTACAGTTGATGAGATATTAGAGCGGTGAC	365
Dp	241	TCAAGCTTAATGTTAGAACATTAACCCATTAAGTACAGTTGATGAGATATTAGAGCGGTGAC	300
QY	366	ATGCACAATACAGACACTCAGACAGATTAAGCCAAAATGATGTGTCAATTAACCTCCAGAG	425
Dp	301	ATGCACAATACAGACACTCAGACAGATTAAGCCAAAATGATGTGTCAATTAACCTCCAGAG	360
QY	426	CTAATATGTCATATGCAAAAAGAGAGAGCGTTAAGACAAAGGACATGAGTCACAAAATTT	485
Dp	361	CTAATATGTCATATGCAAAAAGAGAGAGCGTTAAGACAAAGGACATGAGTCACAAAATTT	420
QY	486	CACATGCGCTTTCGTGCAAAAGAGAGAGCGTTTAATATATCCATGATATTTGCCAAAAG	545
Dp	421	CACATGCGCTTTCGTGCAAAAGAGAGAGCGTTTAATATATCCATGATATTTGCCAAAAG	480
QY	546	AAAGAGAGAAAGAACACACAAATGCTGCGTCAATTATACATATCTGTATGTGCATCATTA	605
Dp	481	AAAGAGAGAAAGAACACACAAATGCTGCGTCAATTATACATATCTGTATGTGCATCATTA	540
QY	606	TTTCAATCAACCTTTTCGTGTACACACACTTCATATATATCAATAGAGTCACTTCAAGTGTGAC	665
Dp	541	TTTCAATCAACCTTTTCGTGTACACACACTTCATATATATCAATAGAGTCACTTCAAGTGTGAC	600

[illegible]

RESULT 8  
US-10-230-365-1  
; Sequence 1, Application US/10230365  
; Publication No. US20030159182A1  
; GENERAL INFORMATION:

Query Match	11.0%;	Score 85.6;	DB 12;	Length 980;
Best Local Similarity	63.1%;	Pred. No. 4.7e-12;		
Matches 149; Conservative	0;	Mismatches 84;	Indels 3;	Gaps 1

QY	541	AAAAAGAAAGGAAAAAGAAACAACAAGCTGGCGTCAATTATATCAATATGTGATGTCAT	600
QY	542	AAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG	601
Db	748	AAAAGCAAAAGAAAGAAAGATATCAAAAGAACATTTGTGATGCCATTTGATATGCTCT	807
QY	601	CATTATTCATCCACCTTTGCTGTACCAACATCTCATATATCATTAAGAGTCACTTCACGCT	660
QY	602	TT	661
Db	808	ATTATATCATCCGCTTTGTGTGTACTTACTTC---TATGTAGTAGTCACTTCATATGT	864
QY	661	GGACATTAACAAACTATCTTAACTTTAGATGCAAGAGCCTTTATCTCACTATTAATG	720
QY	662	GGACATTAACAAACTATCTTAACTTTAGATGCAAGAGCCTTTATCTCACTATTAATG	721
Db	865	GGACATTAACAAACTATCTTAACTTTAGATGCAAGAGCCTTTATCTCACTATTAATGAAG	924
QY	721	CACGATGATTTTCATATGTTTCTCAAAAGAAAGCAATAGCTATATAGTGTCCAAAC	776
QY	722	TT	777
Db	925	GACCAACATATATATCATCTTCTCAACAAAGCAATTAAGTGTGCTCCACAAAAC	980

```

RESULT 9
US-10-076-816-19
: Sequence 19, Application US/10076816
: Publication No. US2003005624A1
: GENERAL INFORMATION:
: APPLICANT: Huang, Ning
: APPLICANT: Rodriguez, Raymond
: APPLICANT: Hagler, Frank E.
: TITLE OF INVENTION: Feed Additive Compositions and Methods

```





; TITLE OF INVENTION: Transgenic Plants  
 ; FILE REFERENCE: 50665-8022, US00  
 ; CURRENT APPLICATION NUMBER: US/10/077,381  
 ; CURRENT FILING DATE: 2002-02-14  
 ; PRIOR APPLICATION NUMBER: US 60/269,199  
 ; PRIOR FILING DATE: 2001-02-14  
 ; PRIOR APPLICATION NUMBER: US 09/847,232  
 ; PRIOR FILING DATE: 2001-05-02  
 ; PRIOR APPLICATION NUMBER: US 60/266,929  
 ; PRIOR FILING DATE: 2001-02-06  
 ; PRIOR APPLICATION NUMBER: US 60/201,182  
 ; PRIOR FILING DATE: 2000-05-02  
 ; NUMBER OF SEQ ID NOS: 33  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 26  
 ; LENGTH: 85  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: GC3 signal peptide sequence  
 US-10-077-381-26

Query Match 6.1%; Score 47.2; DB 14; Length 85;  
 Best Local Similarity 72.6%; Pred. No. 0.012; Indels 0; Gaps 0;  
 Matches 61; Conservative 0; Mismatches 23;

QY 658 TCTGACATTAACAACCTCTATCTTACATTTAGATGCAAGAGCCTTATCTCACTATAA 717  
 DB 2 TGTGGACATTAACAACCTCTATCTTACATCTAGTCATCATCACTTACTTACTATATAA 61  
 QY 718 ATGACAGATGATTTCTCATTTGTT 741  
 DB 62 AAGGACCAACATATCATCATCAT 85

RESULT 13  
 US-10-027-632-212754  
 ; Sequence 212754, Application US/10027632  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, David G.  
 ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
 ; FILE REFERENCE: 108627,129  
 ; CURRENT APPLICATION NUMBER: US/10/027,632  
 ; CURRENT FILING DATE: 2002-04-30  
 ; PRIOR APPLICATION NUMBER: US 60/218,006  
 ; PRIOR FILING DATE: 2000-07-12  
 ; PRIOR APPLICATION NUMBER: US 60/198,676  
 ; PRIOR FILING DATE: 2000-04-20  
 ; PRIOR APPLICATION NUMBER: US 60/193,483  
 ; PRIOR FILING DATE: 2000-03-29  
 ; PRIOR APPLICATION NUMBER: US 60/185,218  
 ; PRIOR FILING DATE: 2000-02-24  
 ; PRIOR APPLICATION NUMBER: US 60/167,363  
 ; PRIOR FILING DATE: 1999-11-23  
 ; PRIOR APPLICATION NUMBER: US 60/156,358  
 ; PRIOR FILING DATE: 1999-09-28  
 ; PRIOR APPLICATION NUMBER: US 60/146,002  
 ; PRIOR FILING DATE: 1999-08-09  
 ; NUMBER OF SEQ ID NOS: 325720  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 212754  
 ; LENGTH: 541  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 US-10-027-632-212754

Query Match 5.4%; Score 41.8; DB 13; Length 541;  
 Best Local Similarity 55.0%; Pred. No. 0.75;  
 Matches 82; Conservative 0; Mismatches 67; Indels 0; Gaps 0;  
 QY 413 TAAACCTCAGAGCTATATGCAATTTGCAAAAAGAGGAGGAGCTTAAAGACAAGGCATG 472

DB 345 TTTAAATGCCACATTCACAGAGTCATTAATGCAATTAATGAGAAACCACTACACAAATTAATG 404  
 QY 473 ACTCAGAAAATTCATCTGCTTGTCTGTCGCAAAAAGAGAGGGCTTATCATTCATGATG 532  
 DB 405 CCAAAAATTAATTAATCTGGAGCTTGTGATTTAAATTAACACTTATCTCATATATACGA 464  
 QY 533 CATATGTCAAAAGAGAGAGAGAAACAA 561  
 DB 465 ATTAAAGGTAAATCAAAATGAGAAATACAA 493

RESULT 14  
 US-10-268-797-3/c  
 ; Sequence 3, Application US/10268797  
 ; Publication No. US20030079250A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Joanne Filiatti  
 ; TITLE OF INVENTION: Nucleic Acid Sequences and Methods of Use for the Production  
 ; FILE REFERENCE: 17138/02/US2/16518,071  
 ; CURRENT APPLICATION NUMBER: US/10/268,797  
 ; CURRENT FILING DATE: 2002-10-11  
 ; PRIOR APPLICATION NUMBER: 09/638,508  
 ; PRIOR FILING DATE: 2000-08-11  
 ; PRIOR APPLICATION NUMBER: US 60/151,224  
 ; PRIOR FILING DATE: 1999-08-26  
 ; PRIOR APPLICATION NUMBER: US 60/172,128  
 ; PRIOR FILING DATE: 1999-12-17  
 ; NUMBER OF SEQ ID NOS: 33  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 3  
 ; LENGTH: 4010  
 ; TYPE: DNA  
 ; ORGANISM: Glycine max  
 US-10-268-797-3

Query Match 5.4%; Score 41.8; DB 14; Length 4010;  
 Best Local Similarity 53.3%; Pred. No. 2;  
 Matches 88; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 351 TATTAGGCGTGACATGACATGACATGACATGACATGACATGACATGACATGACATGACATG 410  
 DB 435 TATTAACAAAACATATATTAAGATCTCAATTTAGTTGTCTAAAGCCAAATGAGAGAG 376  
 QY 411 CATTAACCTCCAGAGCTATATGTCATATTTGCAAAAAGAGAGAGGAGCTTATTAAGCAAGCA 470  
 DB 375 AATGACATTAACAGAAATCTGTGAGATTCATATCATTAACAAAAAAGAGAGAG 316  
 QY 471 TGACTCAGAAAATTTCACTGCTTCCGTCGTCGCAAAAAGAGAGAGG 515  
 DB 315 TAAGAAACCAAAAATTTATTTACCATCATCATGTCGCAAGAAACAAAGAG 271

RESULT 15  
 US-10-268-754-3/c  
 ; Sequence 3, Application US/10268754  
 ; Publication No. US20030084480A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Joanne Filiatti  
 ; TITLE OF INVENTION: Nucleic Acid Sequences and Methods of Use for the Production  
 ; FILE REFERENCE: 17138/02/US2/16518,077  
 ; CURRENT APPLICATION NUMBER: US/10/268,754  
 ; CURRENT FILING DATE: 2002-10-11  
 ; PRIOR APPLICATION NUMBER: 09/638,508  
 ; PRIOR FILING DATE: 2000-08-11  
 ; PRIOR APPLICATION NUMBER: US 60/151,224  
 ; PRIOR FILING DATE: 1999-08-26  
 ; PRIOR APPLICATION NUMBER: US 60/172,128  
 ; PRIOR FILING DATE: 1999-12-17  
 ; NUMBER OF SEQ ID NOS: 33  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 3



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2003, 20:29:05 ; Search time 14935.6 Seconds

(without alignments)  
17056.227 Million cell updates/sec

Title: US-09-847-232a-35

Perfect score: 6227

Sequence: 1 ggtaccatctatcatacatta.....ataatcgcgtgcaagcct 6227

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenBank:\*  
1: gb\_da:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_cm:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sv:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_da:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pin:\*  
35: em\_htg\_rod:\*  
36: em\_htg\_mam:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No	Score	Query Match	Length	DB	ID	Description
1	6227	100.0	6227	6	AX298122	AX298122 Sequence
2	6227	100.0	6227	8	AF395819	AF395819 Oryza sat
3	5692.6	91.4	151479	8	AC093713	AC093713 Oryza sat
4	4658.2	74.8	3651	8	AB021736	AB021736 Oryza sat
5	661	10.6	1065	8	AY072929	AY072929 Oryza sat
6	563.4	9.0	1696	8	RICEB21P	RICEB21P
7	402	6.5	4027	8	HYB121	HYB121
8	266.8	4.3	1493	8	ZMU35063	ZMU35063
9	266.8	4.3	1520	8	MZB0P1	MZB0P1
10	244	3.9	1089	2	MZB0P1	MZB0P1
11	189.2	3.0	108982	2	CNS08C7Q	CNS08C7Q
12	187.6	3.0	183414	8	CNS08C9F	CNS08C9F
13	177.8	2.9	159749	8	AP003020	AP003020 Oryza sat
14	173.4	2.8	152243	2	AC134928	AC134928
15	172.2	2.8	121799	8	OSJN0064	OSJN0064
16	166	2.7	92258	2	AC137997	AC137997 Oryza sat
17	162	2.6	121056	2	AP005510	AP005510 Oryza sat
18	162	2.6	145538	2	AP004875	AP004875 Oryza sat
19	161.6	2.6	174301	2	AP003523	AP003523 Oryza sat
20	160.6	2.6	4442	8	AB053475	AB053475 Oryza sat
21	160.6	2.6	6335	6	BD143755	BD143755 bZIP tran
22	160.6	2.6	6335	6	BD168617	BD168617 bZIP type
23	160.6	2.6	184874	2	AP003800	AP003800 Oryza sat
24	158.8	2.6	157920	2	AP003564	AP003564 Oryza sat
25	158.8	2.6	163279	8	AC078829	AC078829 Oryza sat
26	158.8	2.6	300029	8	AE017105	AE017105 Oryza sat
27	158.4	2.5	130977	2	OSJN00142	OSJN00142
28	158	2.5	144290	2	AC134927	AC134927 Oryza sat
29	156.4	2.5	168267	8	AP005535	AP005535 Oryza sat
30	156.2	2.5	195583	3	AC134234	AC134234 Oryza sat
31	148.4	2.4	147612	2	AP004883	AP004883 Oryza sat
32	148	2.4	134014	8	AC137698	AC137698
33	148	2.4	142248	8	AC107315	AC107315 Genomic s
34	146.8	2.4	144233	8	AP003240	AP003240 Oryza sat
35	146.2	2.3	135712	2	AC108499	AC108499 Oryza sat
36	144.6	2.3	167626	2	OSJN00048	OSJN00048
37	144.2	2.3	141606	2	AP005866	AP005866 Oryza sat
38	144.2	2.3	162221	2	AP005496	AP005496 Oryza sat
39	144	2.3	2000	6	AX656035	AX656035 Sequence
40	143.4	2.3	159193	2	AP004792	AP004792 Oryza sat
41	138.8	2.2	1751	6	BD143754	BD143754 bZIP tran
42	138.8	2.2	1751	6	BD168616	BD168616 bZIP type
43	138.8	2.2	1768	8	AB053472	AB053472 Oryza sat
44	138.4	2.2	115014	2	AC144739	AC144739 Oryza sat
45	137.4	2.2	102008	2	AP004026	AP004026 Oryza sat

# ALIGNMENTS

RESULT 1  
AX298122  
LOCUS AX298122  
DEFINITION Sequence 35 from Patent WO0183792.  
ACCESSION AX298122  
VERSION AX298122.1 GI:17128198  
KEYWORDS  
SOURCE  
ORGANISM Oryza sativa  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophytidae; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Erihartoideae; Oryzaceae; Oryza.  
REFERENCE  
AUTHORS Huang, N., Huang, Y.S., Yang, D. and Schmidt, R.J.  
TITLE Plant transcription factors and enhanced gene expression





Db 4411 GTGATTTTGTATGAAGATGTTGAGCAAGCACTTCCGATATATATGCTAGTTTTGTTGAC 4200  
QY 4201 TGTTCAGAGAGCAATCCAAATCGGAGTCAGCGGCGCTCAGAAAGCAGAAAGCAACT 4260  
Db 4201 TGTTCAGAGAGCAATCCAAATCGGAGTCAGCGGCGCTCAGAAAGCAGAAAGCAACT 4260  
QY 4261 CACTGAATGAGCTGGAGGACAGGTCGATGTTGACATGATTTTGCATATGACAT 4320  
Db 4261 CACTGAATGAGCTGGAGGACAGGTCGATGTTGACATGATTTTGCATATGACAT 4320  
QY 4321 AAAATCTTAATTAAGCTGACTGACTGACTGACTGACTGACTGACTGACTGACTGACT 4380  
Db 4321 AAAATCTTAATTAAGCTGACTGACTGACTGACTGACTGACTGACTGACTGACTGACT 4380  
QY 4381 ATTAAGAGTCGAGAACCTCGCTGTTAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4440  
Db 4381 ATTAAGAGTCGAGAACCTCGCTGTTAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4440  
QY 4441 TGATGCTGCTGTTGACATAGAGTGTAAAGCAGATGTTGAGCCTTGAGCAAGGT 4500  
Db 4441 TGATGCTGCTGTTGACATAGAGTGTAAAGCAGATGTTGAGCCTTGAGCAAGGT 4500  
QY 4501 ATGCTATATAGCTGTTGACATAGAGTGTAAAGCAGATGTTGAGCCTTGAGCAAGGT 4560  
Db 4501 ATGCTATATAGCTGTTGACATAGAGTGTAAAGCAGATGTTGAGCCTTGAGCAAGGT 4560  
QY 4561 CTTTCAAGCTGACTGTTGACATAGAGTGTAAAGCAGATGTTGAGCCTTGAGCAAGGT 4620  
Db 4561 CTTTCAAGCTGACTGTTGACATAGAGTGTAAAGCAGATGTTGAGCCTTGAGCAAGGT 4620  
QY 4621 TATTAATCAACATTTGAGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 4680  
Db 4621 TATTAATCAACATTTGAGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 4680  
QY 4681 AAATGATCAAAATTTGAGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 4740  
Db 4681 AAATGATCAAAATTTGAGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 4740  
QY 4741 ATGAGCGCTGTTGTTCCCGCGCTTGTGATGATGATGATGATGATGATGATGATGATGAT 4800  
Db 4741 ATGAGCGCTGTTGTTCCCGCGCTTGTGATGATGATGATGATGATGATGATGATGATGAT 4800  
QY 4801 TCCCAATCGAAGCAAGCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4860  
Db 4801 TCCCAATCGAAGCAAGCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4860  
QY 4861 TTGCTACTPAACAAGACATGAGAGTAAACAACATGATGATGATGATGATGATGATGATGAT 4920  
Db 4861 TTGCTACTPAACAAGACATGAGAGTAAACAACATGATGATGATGATGATGATGATGATGAT 4920  
QY 4921 GCTCAGAGAGAGAGAGATTTGCTAATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4980  
Db 4921 GCTCAGAGAGAGAGAGATTTGCTAATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4980  
QY 4981 GCTCAGAGAGAGAGAGATTTGCTAATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5040  
Db 4981 GCTCAGAGAGAGAGAGATTTGCTAATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5040  
QY 5041 GCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5100  
Db 5041 GCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5100  
QY 5101 CTGAGAGAGAGAGAGATTTGCTAATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5160  
Db 5101 CTGAGAGAGAGAGAGATTTGCTAATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5160  
QY 5161 ATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5220  
Db 5161 ATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5220  
QY 5221 CTCTCTTAAGGAGCTTGTAGTACCCCTGCTGATCTGCTTAATTAATTAATTAATTAAT 5280  
Db 5221 CTCTCTTAAGGAGCTTGTAGTACCCCTGCTGATCTGCTTAATTAATTAATTAATTAAT 5280

QY 5281 ACCCAATGTTCAATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5340  
Db 5281 ACCCAATGTTCAATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5340  
QY 5341 CTGCTATGCTATCAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5400  
Db 5341 CTGCTATGCTATCAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5400  
QY 5401 GCACCAAGAGCAATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5460  
Db 5401 GCACCAAGAGCAATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5460  
QY 5461 CTCATATCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5520  
Db 5461 CTCATATCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5520  
QY 5521 TGTGATTAAGAGGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5580  
Db 5521 TGTGATTAAGAGGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5580  
QY 5581 ATGCTATTTTCCAGAGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 5640  
Db 5581 ATGCTATTTTCCAGAGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 5640  
QY 5641 GTAAATTTTCTGTTAG 5700  
Db 5641 GTAAATTTTCTGTTAG 5700  
QY 5701 TGAACATTAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5760  
Db 5701 TGAACATTAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5760  
QY 5761 CATTAATCTCGTTATCAAC 5820  
Db 5761 CATTAATCTCGTTATCAAC 5820  
QY 5821 TATGCTGCTTGGCGGAG 5880  
Db 5821 TATGCTGCTTGGCGGAG 5880  
QY 5881 AAGGCATCTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 5940  
Db 5881 AAGGCATCTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 5940  
QY 5941 CTATCTGCTGCTTACAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 6000  
Db 5941 CTATCTGCTGCTTACAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 6000  
QY 6001 CAGATTAGTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6060  
Db 6001 CAGATTAGTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6060  
QY 6061 GAGAGTACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6120  
Db 6061 GAGAGTACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6120  
QY 6121 AAGACATTAATTAAG 6180  
Db 6121 AAGACATTAATTAAG 6180  
QY 6181 ATTAATTAAG 6240  
Db 6181 ATTAATTAAG 6240

RESULT 2  
AF395819 6227 bp DNA linear PLN 03-OCT-2001  
LOCUS Oryza sativa transcription activator REB (Reb) gene, complete cds.  
DEFINITION AF395819  
ACCESSION AF395819  
VERSION AF395819.1 GI:15865781  
KEYWORDS

SOURCE  
ORGANISM  
Oryza sativa (indica cultivar-group)  
Oryza sativa (indica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Eriaristaceae; Oryzaceae; Oryza.

REFERENCE  
AUTHORS  
TITLE  
Yang, D., Wu, L., Hwang, Y.-S., Chen, L. and Huang, N.  
Expression of the REB transcriptional activator in rice grains  
improves the yield of recombinant proteins whose genes are  
controlled by a Reb-responsive promoter  
Proc. Natl. Acad. Sci. U.S.A. 98 (20), 11438-11443 (2001)

JOURNAL  
MEDLINE  
21457336  
PubMed  
11572930  
2 (bases 1 to 6227)  
Yang, D. and Huang, N.  
Direct Submission  
Submitted (27-JUN-2001) Molecular Technology, Applied Phytologics  
Inc., 4110 North Freeway, Sacramento, CA 95854, USA

FEATURES  
SOURCE  
1..6227  
Location/Qualifiers  
/organism="Oryza sativa (indica cultivar-group)"  
/mol\_type="genomic DNA"  
/variety="IR64"  
/sub\_species="Indica"  
/db\_xref="taxon:39946"  
1..>5064  
/gene="Reb"  
1..1006  
/gene="Reb"  
398..727  
/gene="Reb"  
/note="Ta-rich region"  
967..977  
/gene="Reb"  
join(1007..1573,2580..2673,3209..3404,4209..4284,  
4373..4498,4705..>5064)  
/gene="Reb"  
/product="transcription activator REB"  
1007..1573  
/gene="Reb"  
/number=1  
join(1163..1573,2580..2673,3209..3404,4209..4284,  
4373..4498,4705..5064)  
/gene="Reb"  
/function="binds to the upstream activation sequence, UAS,  
and activates the globulin gene promoter"  
/codon\_start=1  
/product="transcription activator REB"  
/protein\_id="AAL10017.1"  
/db\_xref="GI:15865782"  
/translation="MERVSVSEISDPWPVPPPPSAAQAQGGGGVAGSGGGVIA  
GGGGGNNMNCPSWTFKFLLEAVLDSPVNPSPFAEAGGIRGAGVAVDVAKQD  
LSAATTSAAVDPEYNAIKKLEKDLAAVAMRAGTVPPERPGASSILNADVH  
IGAPNSIGNATPVOMLGGSGGSGDLYQNVAVIKPTSSSSRESDDDKRGAA  
ETTGARPADRLQKRSRARSRRKRAHINLEKQVSOLRVNSSLRLRLAD  
VNOKINDAVNRYLKADEVETLRKKMDESVKFMALPPAASDMSLSMPFNS  
PEASTSAAPVIODDPNNPATFNNDIGNNMYMDISSQEDDPVNGALAAKRIK  
PASLDVHSLERHOKRMCGPSGSGTS"  
2580..2673  
/gene="Reb"  
/number=2  
3209..3404  
/gene="Reb"  
/number=3  
4209..4284  
/gene="Reb"  
/number=4  
4373..4498  
/gene="Reb"  
/number=5  
4705..>5064  
/gene="Reb"  
exon  
number=6

I Mendel's Paper

BASE COUNT 1683 a 1312 c 1400 g 1832 t  
ORIGIN  
Query Match 100.0%; Score 6227; DB 8; Length 6227;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 6227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GGTACCATCTATATACATTAATTAACAGAGAGAGAAATGTAATCAATTAATTTT 60  
Db 1 GGTACCATCTATATACATTAATTAACAGAGAGAGAAATGTAATCAATTTT 60  
61 ATGGAGGCTATATTTTATCGGATTTAGTAATTAACGGGCAATTGCTACTTAAGTA 120  
Db 61 ATGGAGGCTATATTTTATCGGATTTAGTAATTAACGGGCAATTGCTACTTAAGTA 120  
121 AAGCTAGCTATGCTATGCTACGCTACGCTACGCTACGCTACGCTACGCTACGCTAC 180  
Db 121 AAGCTAGCTATGCTATGCTACGCTACGCTACGCTACGCTACGCTACGCTACGCTAC 180  
181 TGT 240  
Db 181 TGT 240  
241 GCAAGTTTATAGGATTTCCAGTGTAGCTATGAGGCAATTCATGTCACGGATTTGTAG 300  
Db 241 GCAAGTTTATAGGATTTCCAGTGTAGCTATGAGGCAATTCATGTCACGGATTTGTAG 300  
301 TATATGAAGAAAGAGATCCCAACAAATTAATTAATTTGTATTAAGGAGAAATCGAAT 360  
Db 301 TATATGAAGAAAGAGATCCCAACAAATTAATTAATTTGTATTAAGGAGAAATCGAAT 360  
361 TGAGGTGTATATTCACCAACGAGCTACCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 420  
Db 361 TGAGGTGTATATTCACCAACGAGCTACCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 420  
421 TACGAT 480  
Db 421 TACGAT 480  
481 AT 540  
Db 481 AT 540  
541 TAT 600  
Db 541 TAT 600  
541 TAT 660  
Db 541 TAT 660  
601 AGTATACGAATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 660  
Db 601 AGTATACGAATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 660  
661 TAAAGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 720  
Db 661 TAAAGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 720  
721 TAAAGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 780  
Db 721 TAAAGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 780  
721 TAAAGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 840  
Db 721 TAAAGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 840  
841 TTTTCCCTTAT 900  
Db 841 TTTTCCCTTAT 900  
901 AATTTTCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCA 960  
Db 901 AATTTTCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCA 960  
961 TCCGAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1020  
Db 961 TCCGAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1020

Db 961 TCCGATCATTAATTATTACACCAACAAACAAAATTAGACGACGACGCGCTTCCCAA 1020  
Qy 1021 TCCGATCATTAATTATTACACCAACCAAAACAAAATTAGACGACGACGCGCTTCCCAA 1080  
Db 1021 TCCGATCATTAATTATTACACCAACCAAAACAAAATTAGACGACGACGCGCTTCCCAA 1080  
Qy 1081 AAACCCCTGAAATCCCCCTCGGGAATCCATCGAGAGGTAGCCCGGATCCGCGTGGC 1140  
Db 1081 AAACCCCTGAAATCCCCCTCGGGAATCCATCGAGAGGTAGCCCGGATCCGCGTGGC 1140  
Qy 1141 GAGACGGATTCGATTCGAGATGAGAGGAGTTCCTCGTGAAGAGATCCGACCC 1200  
Db 1141 GAGACGGATTCGATTCGAGATGAGAGGAGTTCCTCGTGAAGAGATCCGACCC 1200  
Qy 1201 ATTTGGGTCCCGCTCCCGCGGAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1260  
Db 1201 ATTTGGGTCCCGCTCCCGCGGAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1260  
Qy 1261 CGGCGTGGCTTCGGGAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1320  
Db 1261 CGGCGTGGCTTCGGGAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1320  
Qy 1321 GAAACGGGTCCCGCGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1380  
Db 1321 GAAACGGGTCCCGCGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1380  
Qy 1381 CCCCGTCCCGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1440  
Db 1381 CCCCGTCCCGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1440  
Qy 1441 GGTGGCGGTTCGATTTAGAGAGCGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1500  
Db 1441 GGTGGCGGTTCGATTTAGAGAGCGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1500  
Qy 1501 GGAACCCCTGAGTACACGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1560  
Db 1501 GGAACCCCTGAGTACACGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1560  
Qy 1561 CGGATGTGGAGGTGAAGCATTCCTCCCGCTGTAGTACTGAGAGCTTCTAGATC 1620  
Db 1561 CGGATGTGGAGGTGAAGCATTCCTCCCGCTGTAGTACTGAGAGCTTCTAGATC 1620  
Qy 1621 GGCATGTGGAGGTGAAGCATTCCTCCCGCTGTAGTACTGAGAGCTTCTAGATC 1680  
Db 1621 GGCATGTGGAGGTGAAGCATTCCTCCCGCTGTAGTACTGAGAGCTTCTAGATC 1680  
Qy 1681 AGACCAATGGCGTCCATGGTTCGATCGCTGCTGTTGAGTGGTGGTGGTGGTGG 1740  
Db 1681 AGACCAATGGCGTCCATGGTTCGATCGCTGCTGTTGAGTGGTGGTGGTGGTGG 1740  
Qy 1741 TCAATGATGATGAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1800  
Db 1741 TCAATGATGATGAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1800  
Qy 1801 ATGCGATGATGATGAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1860  
Db 1801 ATGCGATGATGATGAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1860  
Qy 1861 GCGAGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1920  
Db 1861 GCGAGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1920  
Qy 1921 TGCCTTTTGAATGGAATGATGATGATGATGATGATGATGATGATGATGATGAT 1980  
Db 1921 TGCCTTTTGAATGGAATGATGATGATGATGATGATGATGATGATGATGATGAT 1980  
Qy 1981 ACAAAAAGTTGGGGGATTCACATACCGATGACACTGGCATAAACATCACTCAGTA 2040  
Db 1981 ACAAAAAGTTGGGGGATTCACATACCGATGACACTGGCATAAACATCACTCAGTA 2040  
Qy 2041 TTAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2100  
Db 2041 TTAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2100

Qy 2101 TGAATAATGAGGGGGGACGACGGGGGATTCATAGGCTCCGATGGGTGACAGGGGGA 2160  
Db 2101 TGAATAATGAGGGGGGACGACGGGGGATTCATAGGCTCCGATGGGTGACAGGGGGA 2160  
Qy 2161 CTGAGTCCCGCTGACACCGATGATCCGCTGGATGATGATGATGATGATGATGATG 2220  
Db 2161 CTGAGTCCCGCTGACACCGATGATCCGCTGGATGATGATGATGATGATGATGATG 2220  
Qy 2221 ATTTCTAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2280  
Db 2221 ATTTCTAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2280  
Qy 2281 CATGATTTCTAG 2340  
Db 2281 CATGATTTCTAG 2340  
Qy 2341 TGGCCCTTCATTTGCTGAAATGACATTTGTTGGTGGTGGTGGTGGTGGTGGTGG 2400  
Db 2341 TGGCCCTTCATTTGCTGAAATGACATTTGTTGGTGGTGGTGGTGGTGGTGGTGG 2400  
Qy 2401 TATGTTGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2460  
Db 2401 TATGTTGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2460  
Qy 2461 TCAATTTTGTACATGACTTAAATGACATGACATGACATGACATGACATGACATG 2520  
Db 2461 TCAATTTTGTACATGACTTAAATGACATGACATGACATGACATGACATGACATG 2520  
Qy 2521 AATCATTTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2580  
Db 2521 AATCATTTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2580  
Qy 2581 CCTGCGACAGTTCACCTGAGCTCCGAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2640  
Db 2581 CCTGCGACAGTTCACCTGAGCTCCGAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2640  
Qy 2641 TTTTCAACATGAGCGCTCTAATTCATCGAGAGTATCTATCTGATGATGATGATG 2700  
Db 2641 TTTTCAACATGAGCGCTCTAATTCATCGAGAGTATCTATCTGATGATGATGATG 2700  
Qy 2701 AGATGTTATGAACTACCAATATCTGACATGATGATGATGATGATGATGATGATG 2760  
Db 2701 AGATGTTATGAACTACCAATATCTGACATGATGATGATGATGATGATGATGATG 2760  
Qy 2761 TTTGAAATGAGAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2820  
Db 2761 TTTGAAATGAGAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2820  
Qy 2821 TTAATAGTCCGATTTGATGATGATGATGATGATGATGATGATGATGATGATG 2880  
Db 2821 TTAATAGTCCGATTTGATGATGATGATGATGATGATGATGATGATGATGATG 2880  
Qy 2881 ATAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2940  
Db 2881 ATAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2940  
Qy 2941 GAGCATTTGTTGGCTTGAACATGATGATGATGATGATGATGATGATGATGATG 3000  
Db 2941 GAGCATTTGTTGGCTTGAACATGATGATGATGATGATGATGATGATGATGATG 3000  
Qy 3001 AAAACATTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3060  
Db 3001 AAAACATTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3060  
Qy 3061 CTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3120  
Db 3061 CTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3120  
Qy 3121 GTGCAAAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 3180  
Db 3121 GTGCAAAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 3180



QY	3181	TCCTGTTGGAAATANGGCAATATACAGCAATGCTACTCCAGTTCAAAACATGCTAGT	3240
Db	3181	CTCTGTTGGAAATANGGCAATATACAGCAATGCTACTCCAGTTCAAAACATGCTAGT	3240
QY	3241	GGCCCAAGTGGGGGAGATCGGGGCTCAACTTGTGACAGATGTGATGCTTTAAAGAG	3300
Db	3241	GGCCCAAGTGGGGGAGATCGGGCTCAACTTGTGACAGATGTGATGCTTTAAAGAG	3300
QY	3301	CCCAACAGCTCTTCATCAAGGGAGCACTGATGATGATGACATGAAGAGAAAGCTGAG	3360
Db	3301	CCCAACAGCTCTTCATCAAGGGAGCACTGATGATGATGACATGAAGAGAAAGCTGAG	3360
QY	3361	ACCACTGGAAACAGCAAGACTGCTGATCAAGAATTTACAGCAAGAGAGATTCATCTGCT	3420
Db	3361	ACCACTGGAAACAGCAAGACTGCTGATCAAGAATTTACAGCAAGAGAGATTCATCTGCT	3420
QY	3421	TCCTTGTAAATATAGATTCTGTACATTAATTAACCTACCTGACATGACATGATGCTCA	3480
Db	3421	TCCTTGTAAATATAGATTCTGTACATTAATTAACCTACCTGACATGACATGATGCTCA	3480
QY	3481	TTTTACACTTAGCCCTTCAGTGGATTTCCTCACTTCATCCGGTAGCCCTTACAGTTCCT	3540
Db	3481	TTTTACACTTAGCCCTTCAGTGGATTTCCTCACTTCATCCGGTAGCCCTTACAGTTCCT	3540
QY	3541	ATTGCATGCAATATAGATCTTTTACCTACCAATTAATTAAGTCTCTGTGTGCATTACTCAGT	3600
Db	3541	ATTGCATGCAATATAGATCTTTTACCTACCAATTAATTAAGTCTCTGTGTGCATTACTCAGT	3600
QY	3601	GCTTAGTGCTCTGAGAGAGAGAGAAATTTGTATAGGCGCTATACAGTAGCACTTGTCTGC	3660
Db	3601	GCTTAGTGCTCTGAGAGAGAGAGAAATTTGTATAGGCGCTATACAGTAGCACTTGTCTGC	3660
QY	3661	TACTGTTTATATACATTAAGCAATTTGGATGATTAATTAATGTGAGTCAATCAATATACC	3720
Db	3661	TACTGTTTATATACATTAAGCAATTTGGATGATTAATTAATGTGAGTCAATCAATATACC	3720
QY	3721	TTATGTCAAGAATTAAGTATATACCTCAAGAGCTTTACCTTCATTAAGCAATTAACC	3780
Db	3721	TTATGTCAAGAATTAAGTATATACCTCAAGAGCTTTACCTTCATTAAGCAATTAACC	3780
QY	3781	GGATTTGCAATCAAGTATATACCTCAAGAGCTTTACCTTCATTAAGCAATTAACC	3840
Db	3781	GGATTTGCAATCAAGTATATACCTCAAGAGCTTTACCTTCATTAAGCAATTAACC	3840
QY	3841	CAATTAACCAAGGATTTGATACATTAATGACTATGATTACCAAGATTCATTTGGCACTAT	3900
Db	3841	CAATTAACCAAGGATTTGATACATTAATGACTATGATTACCAAGATTCATTTGGCACTAT	3900
QY	3901	TTTTCCAAATTAAGTTGGATATAGTCTCGAGTGGCTGTAAAAAATAGTCAGAGAGGT	3960
Db	3901	TTTTCCAAATTAAGTTGGATATAGTCTCGAGTGGCTGTAAAAAATAGTCAGAGAGGT	3960
QY	3961	ACATGATATGTACAGCATACAAAGTATGGGGATTTTGATATGACAGTGTACACCC	4020
Db	3961	ACATGATATGTACAGCATACAAAGTATGGGGATTTTGATATGACAGTGTACACCC	4020
QY	4021	ACATTTGCTCACTAAATATAAATATTCACAGCTCAACGATGATATGGTGATTCATTT	4080
Db	4021	ACATTTGCTCACTAAATATAAATATTCACAGCTCAACGATGATATGGTGATTCATTT	4080
QY	4081	ATACTTGTATGTTTATTAATGTCTCTGGTCAGACATTAATTAAGGCTGCTCTTGG	4140
Db	4081	ATACTTGTATGTTTATTAATGTCTCTGGTCAGACATTAATTAAGGCTGCTCTTGG	4140
QY	4141	GTGATTTTGTATAGATTTGTAGCAAGCACTTCTGATATATATGCTAGTTTGTAGCC	4200
Db	4141	GTGATTTTGTATAGATTTGTAGCAAGCACTTCTGATATATATGCTAGTTTGTAGCC	4200
QY	4201	TTTTCCAGGAAGAAATCAATCGAGAGCTGACGCAAGGCGCTCAAGAAAGAAAGAGAGCT	4260
Db	4201	TTTTCCAGGAAGAAATCAATCGAGAGCTGACGCAAGGCGCTCAAGAAAGAAAGAGAGCT	4260
QY	4261	CACWTGGAATGAGCTGAGGCAACAGGTGTGATTTACATTAATTAATTTTGATTAACAT	4320

Db	4261	CACCTGATAGTGGAGGACACAGTGTGATGTTGCATATGTTATTTTGATTAAGACAT	4320
OY	4321	AAAATCCTAATATACAGCTACTGACTTCAGTTATGATGATTACTTGTTCAGGATATCCGA	4380
Db	4321	AAAATCCTAATATACAGCTACTGACTTCAGTTATGATGATTACTTGTTCAGGATATCCGA	4380
OY	4381	ATTAAAGTGTGAGACTCCTCGCTGTAAAGGCTCTTGCTATGTTTAACCAAGTATACA	4440
Db	4381	ATTAAAGTGTGAGACTCCTCGCTGTAAAGGCTCTTGCTATGTTTAACCAAGTATACA	4440
OY	4441	TGATGCGCTGTGCAATATAGTGTAAAGACAGTGTGACCTTGAGACCAAGT	4500
Db	4441	TGATGCGCTGTGCAATATAGTGTAAAGACAGTGTGACCTTGAGACCAAGT	4500
OY	4501	ATGCTATATATGCTTTTGCAATATGCAATCCATGGAATTGCTACTTTGGCTTGTTCAAA	4560
Db	4501	ATGCTATATATGCTTTTGCAATATGCAATCCATGGAATTGCTACTTTGGCTTGTTCAAA	4560
OY	4561	CTTTCACAGGACTTGTGTACCGCTGTATATAGAAAGATATATCCGCTACATATATAC	4620
Db	4561	CTTTCACAGGACTTGTGTATCCCTGTATATTAAGAAATATATCCGCTACATATATAC	4620
OY	4621	TATATAATACCATTTGGCCAGTCCAAACATGATTATTAATAGGTAAATCTGACATTTG	4680
Db	4621	TATATAATACCATTTGGCCAGTCCAAACATGATTATTAATAGGTAAATCTGACATTTG	4680
OY	4681	AAATGTATCAAAATATCGAGGTGAAATAGGACGAGGAACTGGTGAAGCGGGTACAGGC	4740
Db	4681	AAATGTATCAAAATATCGAGGTGAAATAGGACGAGGAACTGGTGAAGCGGGTACAGGC	4740
OY	4741	ATGAACCGGTTGTTCCCGCCGCTTGTGATATGTCAATCCCTACAGATCCCATTAACAC	4800
Db	4741	ATGAACCGGTTGTTCCCGCCGCTTGTGATATGTCAATCCCTACAGATCCCATTAACAC	4800
OY	4801	TCCCCATCTATAGACACATGTCAGAGCTGCTGTTCCATATCAAGATGACCCGAAATAC	4860
Db	4801	TCCCCATCTATAGACACATGTCAGAGCTGCTGTTCCATATCAAGATGACCCGAAATAC	4860
OY	4861	TTCGCTACTTAACACACATCGAGAGGTAAACAACTACATCCCGACATACCTTTTG	4920
Db	4861	TTCGCTACTTAACACACATCGAGAGGTAAACAACTACATCCCGACATACCTTTTG	4920
OY	4921	GCTCAGAGGACAGACTTGTCAATGCGCTGTGCTGCGGACAGATTGGCCGGCA	4980
Db	4921	GCTCAGAGGACAGACTTGTCAATGCGCTGTGCTGCGGACAGATTGGCCGGCA	4980
OY	4981	GCTCGCGACAGCGGCTGCGAGCTCGAGCACTCCCAAAAGAGATGTGCTGGGCG	5040
Db	4981	GCTCGCGACAGCGGCTGCGAGCTCGAGCACTCCCAAAAGAGATGTGCTGGGCG	5040
OY	5041	GCTTCGTCGGGTGACAGTCTCGAGACCGAAATCCAGACTCTCTTGCTGTAAAGACA	5100
Db	5041	GCTTCGTCGGGTGACAGTCTCGAGACCGAAATCCAGACTCTCTTGCTGTAAAGACA	5100
OY	5101	CTGCGAGAGGAATATATGATGGAACAGCGGTGAATGCTAATCGTAGAGATTGATG	5160
Db	5101	CTGCGAGAGGAATATATGATGGAACAGCGGTGAATGCTAATCGTAGAGATTGATG	5160
OY	5161	ATTGTTGGTCTGCTGTGCTGATATGTGCATTTTGTAGGAGACACCTTTATGACCT	5220
Db	5161	ATTGTTGGTCTGCTGTGCTGATATGTGCATTTTGTAGGAGACACCTTTATGACCT	5220
OY	5221	CTTCTCTTAAGGACTTATAGACCTCTGTGATCTATGCTCTTAATATATATACAT	5280
Db	5221	CTTCTCTTAAGGACTTATAGACCTCTGTGTGATCTATGCTCTTAATATATATACAT	5280
OY	5281	AACCAAAATGTCAATGGTGTGATGCGCTGCTCCCTAATTTGAACACATGATTCAGGAC	5340
Db	5281	AACCAAAATGTCAATGGTGTGATGCGCTGCTCCCTAATTTGAACACATGATTCAGGAC	5340
OY	5341	CTGCTATGCTATCAATTAATATTTGATGCACTCTCTCTGTCTTTGCTCTTAA	5400
Db	5341	CTGCTATGCTATCAATTAATATTTGATGCACTCTCTCTGTCTTTGCTCTTAA	5400

```

Db      5341  CAGTATGCTATCATCATATATATTTGATGATGATGCTTCTCTTCTTTGCTTAA 5400
QY      5401  GCACCAAGCCTAAAGATACACCTTTGACCTGTCTATTGATGCAAGGCTAAG 5460
Db      5401  GCACCAAGCCTAAAGATACACCTTTGACCTGTCTATTGATGCAAGGCTAAG 5460
QY      5461  CTCATATCTAGGCTGCTGATGATGATGATGATGATGATGATGATGATGATG 5520
Db      5461  CTCATATCTAGGCTGCTGATGATGATGATGATGATGATGATGATGATGATG 5520
QY      5521  TGGTGAATGAAGGAGCTGTACATCCCTAATACATGCTATTTGATGATGATG 5580
Db      5521  TGGTGAATGAAGGAGCTGTACATCCCTAATACATGCTATTTGATGATGATG 5580
QY      5581  ATGCAATTTTACCCAGATCCATCCAGTTCATACATGATGATGATGATGATG 5640
Db      5581  ATGCAATTTTACCCAGATCCATCCAGTTCATACATGATGATGATGATGATG 5640
QY      5641  GTAAATTTTGTGTTAGAGAAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 5700
Db      5641  GTAAATTTTGTGTTAGAGAAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 5700
QY      5701  TGAACAAATATCATTGATGATGATGATGATGATGATGATGATGATGATGAT 5760
Db      5701  TGAACAAATATCATTGATGATGATGATGATGATGATGATGATGATGATGAT 5760
QY      5761  CATTAATCTCCGTATACCAACCAACCAACCAACCAACCAACCAACCAACCA 5820
Db      5761  CATTAATCTCCGTATACCAACCAACCAACCAACCAACCAACCAACCAACCA 5820
QY      5821  TCATGCTTTGCGGGGGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 5880
Db      5821  TCATGCTTTGCGGGGGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 5880
QY      5881  AAGGCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCC 5940
Db      5881  AAGGCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCC 5940
QY      5941  CTATCTTCTGCTCTCAAGATGATGATGATGATGATGATGATGATGATGAT 6000
Db      5941  CTATCTTCTGCTCTCAAGATGATGATGATGATGATGATGATGATGATGAT 6000
QY      6001  CAGATATGCTATCATCATATATATTTGATGATGATGATGATGATGATGATG 6060
Db      6001  CAGATATGCTATCATCATATATATTTGATGATGATGATGATGATGATGATG 6060
QY      6061  GAGTGAACATGCGCTGCGCCGTTTCTCTACCGAAACATGCGCAATGAAAG 6120
Db      6061  GAGTGAACATGCGCTGCGCCGTTTCTCTACCGAAACATGCGCAATGAAAG 6120
QY      6121  AAGCAATATATAGGCAATGAGGATGAGGATGAGGATGAGGATGAGGATGAG 6180
Db      6121  AAGCAATATATAGGCAATGAGGATGAGGATGAGGATGAGGATGAGGATGAG 6180
QY      6181  ATTAATATCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 6227
Db      6181  ATTAATATCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 6227

```

RESULT 3  
 AC093713  
 LOCUS Oryza sativa chromosome 3 BAC OSJNB0094F01 genomic sequence,  
 DEFINITION complete sequence.  
 AC093713 151479 bp DNA linear JUN 29-MAY 2003  
 AC093713  
 VERSION  
 KEYWORDS  
 ORGANISM  
 Oryza sativa (japonica cultivar-group)  
 Oryza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Eriophytidae; Oryzaceae; Oryza.  
 REFERENCE 1 (bases 1 to 151479)

AUTHORS  
 Buell, C.R., Yuan, Q., Ouyang, S., Liu, J., Gansberger, K., Jones, K.M.,  
 Overton, II, L.L., Tsirlin, T., Kim, M.M., Bera, J.J., Jin, S.S.,  
 Vadenosh, D.W., Tallon, L.J., Koo, H., Zismann, Y., Heide, J., Blunt, S.,  
 Vanaken, S.S., Riedmiller, S.B., Utterback, T.T., Feldlyum, T.V.,  
 Yang, Q.Q., Haas, B.J., Suh, B.B., Peterson, J.J., Quackenbush, D.,  
 White, O., Salzberg, S.L., and Fraser, C.M.  
 Oryza sativa chromosome 3 BAC OSJNB0094F01 genomic sequence  
 Unpublished  
 2 (bases 1 to 151479)  
 Buell, R.  
 Direct Submission  
 Submitted (08-SEP-2001) The Institute for Genomic Research, 9712  
 Medical Center Dr., Rockville, MD 20850, USA  
 3 (bases 1 to 151479)  
 Buell, R.  
 Direct Submission  
 Submitted (29-MAY-2003) The Institute for Genomic Research, 9712  
 Medical Center Dr., Rockville, MD 20850, USA, rbuell@tigr.org  
 On May 29, 2003 this sequence version replaced gi:15092997.  
 Address all correspondence to: rice@tigr.org

COMMENT  
 BAC clone OSJNB0094F01 is from Oryza sativa chromosome 3  
 The orientation of the sequence is from SP6 to T7 end of the BAC  
 clone.  
 Genes were identified by a combination of several methods: Gene  
 prediction programs including Jgenesh (<http://www.softberry.com/>),  
 GenScan and GenScan+ (Chris Burge,  
<http://CCR-081.mt.edu/GENSCAN.html>), GenMarkES (Mark Borodovsky,  
<http://genemark.biology.gatech.edu/Genemark/>), and GeneSplicer  
 (Michael Pertea and Steven Salzberg, contact [mpertea@tigr.org](mailto:mpertea@tigr.org)),  
 searches of the complete sequence against a peptide database and  
 the plant EST database at TIGR (<http://www.tigr.org/db/tgsl.shtml>).  
 Annotated genes are named to indicate the level of evidence for  
 their annotation. Genes with similarity to other proteins are named  
 after the database hits. Genes without significant peptide  
 similarity but with EST similarity are named as unknown proteins.  
 Genes without protein or EST similarity, that are predicted by more  
 than two gene prediction programs over most of their length are  
 annotated as hypothetical proteins. Genes encoding tRNAs are  
 predicted by tRNAscan-SE (Sean Eddy,  
<http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats are  
 identified by RepeatMasker (Arian Smit,  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>).  
 This BAC overlaps with rice BAC OSJNB0021G19 (AC092076) and  
 OSJNB0052F07 (AC104321).

FEATURES  
 source  
 Location/Qualifiers  
 1..151479  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /mol\_type="genomic DNA"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:39947"  
 /chromosome="3"  
 /clone="OSJNB0094F01"  
 /note="japonica cultivar-group"  
 818..3244  
 /gene="OSJNB0094F01.1"  
 /note="highly similar to 40S ribosomal protein S15  
 GB:AK97632 GI:15425963 (Elaeis oleifera); EST  
 AU057418, B130629 from this gene"  
 join(818..888,997..1151,1236..1293,2816..3344)  
 /gene="OSJNB0094F01.1"  
 join(886..888,997..1151,1236..1293,2816..3064)  
 /gene="OSJNB0094F01.1"  
 /codon\_start=1  
 /product="putative 40S ribosomal protein S15"  
 /protein\_id="AAP4465.1"  
 /db\_xref="GI:31126745"  
 /translation="MADVEAEVAAAGAPKRTFRKSYSGVDLALDMSDIDVQ  
 LFPARARFQGLIKRPMALIKLRKADAPAGKPEPVRILRMILIVPEMISI  
 VGVNKGTFNVEIKPEMIGHYLAEEFISYKPKHGRRGIGATSSRFIDK"  
 complement(5616..7143)  
 /gene="OSJNB0094F01.2"

gene

	Query Match	Best Local Similarity	Matches	Conservative	0;	Mismatches	89;	Indels	108;	Gaps	22;
Yy	1	GSTRACCATCTAATACATTAAATACAAAGAGAAATGCAATATGCAATATTTATTTT	60								
Db	101367	GGTACCCATCTAATACATTAAATACAAAGAGAAATGCAATATTTATTTT	101426								
Yy	61	ATGGAGGCGATA-TTTTATCGAGTTTAAATATACGGGCAATTCGCTACTAGT	119								
Db	101427	ATGGAGGCGATA-TTTTATCGAGTTTAAATATACGGGCAATTCGCTACTAGT	101486								
Yy	120	AAAGCTACGATGACATATGCGCTACCGCTACCGTGTGGAATTTGCGATAGCAT	179								
Db	101487	AAAGCTACGATGACATATGCGCTACCGCTACCGTGTGGAATTTGCGATAGCAT	101546								
Yy	180	CTGTGCTGTTGGCAGTTAGGCTACTTGATTAACCTCCACCGCTGAAACGAGGGGTTT	239								
Db	101547	CTGTGCTGTTGGCAGTTAGGCTACTTGATTAACCTCCACCGCTGAAACGAGGGGTTT	101606								
Yy	240	CGCAGGTTTAAAGATTGCGCAAGTTTAAGCTAGGCGCAATTCATGTTACGCGTATTGTGA	299								
Db	101607	CGCAGGTTTAAAGATTGCGCAAGTTTAAGCTAGGCGCAATTCATGTTACGCGTATTGTGA	101666								
Yy	300	GTAATGTGAAAGAGAGATGCCCAACAATTTAATTTGATTAAGGAGAGAAATCGAAC	359								
Db	101667	GTAATGTGAAAGAGAGATGCCCAACAATTTAATTTGATTAAGGAGAGAAATCGAAC	101728								
Yy	360	TTGAGGTGCTAATTAATCAACACGAGCTACTCCCTCGGTTTCATATATATATACATAT	419								

D	101727	TTGAGGGTGTCTAATTCACCAACCGAGCTACCTCCCTCCGTTTAATATACGTAATATACAT	101786
Q	420	ATACGTAATATATACGTAATATACATATACGTAATATACATATAT	466
D	101787	ACACGTAATATATACATATATAGTAATATACGTAATATATATATATATACGTAATATACACAC	101846
Q	467	AT	526
D	101847	ACACACAT	101906
Q	527	AATGCTAAAAAGTTTAT	586
D	101907	AATGCTAAAAAGTTTAT	101953
Q	567	TTCTGCACTGTGTAGTATATACGATATACGAAATATATATATATATATATATATATATATATATAT	646
D	101954	TTCTGCACTGTGTAGTAT	102013
Q	647	CGCGCGAGGAAAT	706
D	102014	CGCGCGAGGAAAT	102061
Q	707	TTCCACTGAT	766
D	102062	-----GAT	102114
Q	767	TATACACACAGTAT	826
D	102115	TATACACACAGTAT	102174
Q	827	GTTACAAACGCAATATTTTGGCGCTTGTATATATATATATATATATATATATATATATATATATAT	886
D	102175	GTTACAAACGCAATATTTTGGCGCTTGTATATATATATATATATATATATATATATATATATATAT	102225
Q	887	TGCTAT	946
D	102226	TAT	102285
Q	947	CGTATACAGGACAT	1006
D	102286	CGTATACAGGACAT	102345
Q	1007	AGCGCGCTTCCCAATCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT	1066
D	102346	AGCGCGCTTCCCAATCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT	102393
Q	1067	GCTTTCTCCGATATCAAACCTCCGATATCCCTCCGATATCCCTCCGATATCCCTCCGATATCCCTCCG	1126
D	102394	GCTTTCTCCGATATCAAACCTCCGATATCCCTCCGATATCCCTCCGATATCCCTCCGATATCCCTCCG	102453
Q	1127	CGATCCGCTCGGCGAGACGGGATTCGATTCGCGATTCGCGAGACGGGATTCCTCCGTGGAG	1186
D	102454	CGATCCGCTCGGCGAGACGGGATTCGATTCGCGATTCGCGAGACGGGATTCCTCCGTGGAG	102513
Q	1187	GAGATCTCCGACCCATCTTCTGGGTCGCGCGCTCCGCGCGCATTCGCGCGCGCGCGCGCGCGCG	1246
D	102514	GAGATCTCCGACCCATCTTCTGGGTCGCGCGCTCCGCGCGCATTCGCGCGCGCGCGCGCGCGCG	102573
Q	1247	CAGCAGGCGCGCGCGCGCTTCGCGAGGATGCTGATACGCGGCGCGCGCGCGCGCGCGCGCGCG	1306
D	102574	CAGCAGGCGCGCGCGCGCTTCGCGAGGATGCTGATACGCGGCGCGCGCGCGCGCGCGCGCGCGCG	102633
Q	1307	GGCGGGAACGCAATACACGGTCCCTCGTGGAGTGGTACTTCCAGAGTTCTTGGAGGAG	1366
D	102634	GGCGGGAACGCAATACACGGTCCCTCGTGGAGTGGTACTTCCAGAGTTCTTGGAGGAG	102693
Q	1367	GGCGGCTCGATAGACCCCGCTCCGAAACCTCCGAGACCCCGAGGCGGAGAGGATCAG	1426
D	102694	GGCGGCTCGATAGACCCCGCTCCGAAACCTCCGAGACCCCGAGGCGGAGAGGATCAG	102753
Q	1427	GGCGGAGAGGGGTGTGTCCGCTCATGTATAGACACCGCACTT-----CG	1474
D	102754	GGCGGAGAGGGGTGTGTCCGCTCATGTATAGACACCGCACTT-----CG	102813

QY	1475	GC	CGCGGCGAGCAGCAGACCGCGGTGTGAGACCCCTTGAGATCAACAGCGATGCTGAAGCAG	1534
Db	102814	GC	CGCGCGGCGAGCAGCAGACCGCGGTGTGAGACCCCTTGAGATCAACAGCGATGCTGAAGCAG	102873
QY	1535	AAG	CTGGAGAAAGACACCGCGCGGGTGGCCATGCTGGAGGGGAGACGACCATTCG-----CCG	1589
Db	102874	AAG	CTGGAGAAAGACACCGCGCGGGTGGCCATGCTGGAGGGGAGACGACCATTCG-----CCG	102933
QY	1590	CCC	CTCTACTACTGAGAGGCTTACAGAGATGGCAATGGCAATGGTACACCTGTTTGACATGAAAT	1649
Db	102994	CCC	CTCTACTACTGAGAGGCTTACAGAGATGGCAATGGCAATGGTACACCTGTTTGACATGAAAT	102993
QY	1650	GTT	TATATAGATATTAGATCGGGCATTTCTATATAGCAAAAGCGTCACATGGCTCTGCATAG	1709
Db	102994	GTT	TATATAGATATTAGATCGGGCATTTCTATATAGCAAAAGCGTCACATGGCTCTGCATAG	103053
QY	1710	CG	CTGCTGTGAGTACGCGGTGTGTGGTTGCACTATAGTATGATGAGGTTTGCGATATGAC	1769
Db	103054	CG	CTGCTGTGAGTACGCGGTGTGTGGTTGCACTATAGTATGATGAGGTTTGCGATATGAC	103107
QY	1770	AA	GGGAAAGCTCTCAATACCTCGGATATGAGATGGAATGGGATATGGAATGGAACCT--GGGAAAT	1828
Db	103108	AA	GGGAAAGCTCTCAATACCTCGGATATGAGATGGAATGGGATATGGAATGGAACCT--GGGAAAT	103167
QY	1829	GG	CAGTATGATACAGTCAATATCTTCTACTAACTGCGAGACAGTGGTTTGTCATTTGCA	1888
Db	103168	GG	CAGTATGATATAGTCAATATCTTCTACTAACTGCGAGACAGTGGTTTGTCATTTGCA	103227
QY	1889	AAT	ATATTTTGTATGAGGGCGCTGTAACTGT--CATTGCCCTTTTATGATTTGCGAATATGTAC	1947
Db	103228	AAT	ATATTTTGTATGAGGGCGCTGTAACTGTAACTGTAACTGCTTTTATGATTTGCGAATATGTAC	103287
QY	1948	TTT	TATGCAAGATTTTATGATTTGGCGGATCCAGAAACAAAAGTTGGGGGAGATTAACAATAC	2007
Db	103288	TTT	TATGCAAGATTTTATGATTTGGGGGATCCAGAAACAAAAGTTGGGGGAGATTAACAATAC	103347
QY	2008	CG	AGTACACTGCGATAAACAATCATCATCTCAGTATTTAACTATGTAAATGCTAAATGCTATTAGA	2067
Db	103348	CG	AGTACACTGCGATAAACAATCATCATCTCAGTATTTAACTATGTAAATGCTAAATGCTATTAGA	103407
QY	2068	GAC	CTTATACACTCTCTTCTCTTCTTATACACATGAGTGAAGAAAATATGAAAGGGGGAGCTAGGG	2127
Db	103408	GAC	CTTATATACACTCTCTTCTCTTCTTATACACATGAGTGAAGAAAATATGAAAGGGGGAGCTAGGG	103467
QY	2128	GG	GATATCCAAAGGGTCCGATGGGTGCAAGGGGGGAC--TGAGTCCGCCCTGCGACCCACGTTGA	2186
Db	103468	GG	GATATCCAAAGGGTCCGATGGGTGCAAGGGGGGACCTTGAGTCCGCCCTGCGACCCACGTTGA	103527
QY	2187	AT	CGCCCTGCGAGTACGCTATATACGTGTACAGGACCATTTCTAAGTGTCTTGCTTAATTGG	2246
Db	103528	AT	CGCCCTGCGAGTACGCTATATACGTGTACAGGACCATTTCTAAGTGTCTTGCTTAATTGG	103587
QY	2247	GT	GATGATCGAGCTTATTTTGTCTTTTCTATATGCGTACATCATTTCTAAGAAAGAAAT	2306
Db	103588	GT	GATGATCGAGCTTATTTTGTCTTTCTATATGCGTACATCATTTCTAAGAAAGAAAT	103647
QY	2307	AG	CCCTATATATGTGCTCCAGAAATTTGAGATACCTCGGCCCTTCATTTGCTGAATTTAGCC	2366
Db	103648	AG	CCCTATATATATGTGCTCCAGAAATTTGAGATACCTCGGCCCTTCATTTGCTGAATTTAGCC	103706
QY	2367	TAT	TATGTTGGTGGCCCTGAGTTTCTTCTCCAGCTATATGTTGGCAATGTTGGCATGTTGGCATGCG	2426
Db	103707	TAT	TATGTTGGTGGTGG--CCCTGAGTTTCTCCAGCTTATGTTGGCAATGTTGGCATGTTGGCATGCG	103765
QY	2427	CT	CTGCTTTTGGCCCTATATATTTTATATTTTGGCAATCATTTTGTG-----ACATGACTTAA	2481
Db	103766	CT	CTGCTTTTGGTGG--CCATATATTTATTTATTTTGGCAATCATTTTGTGCAATTTGCAATTTGCAATTA	103824
QY	2482	AAT	G--ACACTGAGACAACTGACGACGTATGG--TTATCTTATATCAATTTATGTATTTCT	2538
Db	103825	AAT	GAACTCTGAGACAACTGACGACGTATGGTTTATCTTATATCAATTTATGTATTTCT	103884

QY	2538	GTTCATTTTATCGCTAGCTATGTCATTTTCAATTCAAGGCTTCGSCACAGTTCCAC	2598
Db	103885	GTTCATTTTATCAATGCTAGCTATGTCATTTTCAATTCAAGGCTTCGSCACAGTTCCAC	103944
QY	2599	CTGACCGTCGCGAGCGGGTTCATGCTGCGGAATGAGATGTTTCAACAATPAGGCGCTG	2658
Db	103945	CTGACCGTCGCGAGCGGGTTCATGCTGCGGAATGAGATGTTTCAACAATPAGGCGCTG	104004
QY	2659	CTAATTCATCGAGAGGACTATATCTTATCTGTTACATTTTCAGATGTTTATGAAACTAC	2718
Db	104005	CTAATTCATCGAGAGGACTATATCTTATCTGTTACATTTTCAGATGTTTATGAAACTAC	104064
QY	2719	CGAAATATCGGACAAATTCAGATGCGGATTAATTTTGGTTCTGTTGAAATGAGTGCAG	2778
Db	104065	CGAAATATCGGACAAATTCAGATGCGGATTAATTTTGGTTCTGTTGAAATGAGTGCAG	104124
QY	2779	TGTATTTGCTGTCACGTCAATCAAAATAGTTCGAAAGCATTAATAATAGTTCGCAATT	2838
Db	104125	TGTATTTGCTGTCACGTCAATCAAAATAGTTCGAAAGCATTAATAATAGTTCGCAATT	104184
QY	2839	GTTATGATTTTGTGAACATPAGAAATGTTTGTCTTAAGTATGATGGTTATGTTTGT	2898
Db	104185	GTTATGATTTTGTGAACATGTTGAATTTGTATAGCTTAAGTATGATGGTTATGTTTGT	104244
QY	2899	GGAGTCCCTTAATCACTTTCATTCCTGACGCGACGTCGACAGACATTTGTTGTTGCC	2958
Db	104245	GGAGTCCCTTAATCACTTTCATTCCTGACGCGACGTCGACAGACATTTGTTGTTGCC	104304
QY	2959	TTGACCATGAATGAAGACCTCCCTGTTCTGTATGATGCTCACAGAAAAATATTTGATTTA	3018
Db	104305	TTGACCATGAATGAAGACCTCCCTGTTCTGTATGATGCTCACAGAAAAATATTTGATTTA	104364
QY	3019	TGCACCTTGAATCCTTATGAGTCTGTGCAAGATGAGGCACTAGCTTATGAAATGAGTGTGA	3078
Db	104365	TGCACCTTGAATCCTTATGAGTCTGTGCAAGATGAGGCACTAGCTTATGAAATGAGTGTGA	104424
QY	3079	CTTAAATAGCTGTTGTTATCATGATGATTTGTCTGTATGTAATGTCGACAAAACAGGATG	3138
Db	104425	CTTAAATAGCTGTTGTTATCATGATGATTTGTCTGTATGTAATGTCGACAAAACAGGATG	104484
QY	3139	CTAGCTTTGATTTGATATTCATCAAGCCGCGTTTACTTATGTCGTTGGTGGACAAAGGG	3198
Db	104485	CTAGCTTTGATTTGATATTCATCAAGCCGCGTTTACTTATGTCGTTGGTGGACAAAGGG	104544
QY	3199	CACATATCAGGCAATGCTPACTCGAGTTCAAAACATGCTAATAGTGGCCCAAGTGGGGAATG	3258
Db	104545	CACATATCAGGCAATGCTPACTCGAGTTCAAAACATGCTAATAGTGGCCCAAGTGGGGAATG	104604
QY	3259	GGCTACAGTTGGTATGACGAATGTTATGTCCTGTAAAGAGGCCACAGCTCTTCATCA	3318
Db	104605	GGCTACAGTTGGTATGACGAATGTTATGTCCTGTAAAGAGGCCACAGCTCTTCATCA	104664
QY	3319	AGGGAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	3378
Db	104665	AGGGAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	104724
QY	3379	CTGCTGATCAAAATATACAGAGAGATGATCATTCATGCTCTCTCTGTAATATGATTC	3438
Db	104725	CTGCTGATCAAAATATACAGAGAGATGATCATTCATGCTCTCTCTGTAATATGATTC	104784
QY	3439	TGTACATTAATTAACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTAC	3498
Db	104785	TGTACATTAATTAACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTAC	104844
QY	3499	CAGTTGATTTCCACTTTCAATCCGGATGAGCTTTCAAGTTTCCATTCGATGCAATATATGA	3558
Db	104845	CAGTTGATTTCCACTTTCAATCCGGATGAGCTTTCAAGTTTCCATTCGATGCAATATATGA	104904
QY	3559	TCTTTTACCTACCATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	3618
Db	104905	TCTTTTACCTACCATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	104964
QY	3619	GAGAGAAATTTGATGGAATTTACAGATGACGACTTGTCTCTACTTGTATTATGACATA	3678

Db	104965	GAGAGGATTGTAAGGCCATTACACGAGACCTTTGGCTCTCACTGTTTATTTGACATA	105024
QY	3679	AGCAATTTGGGATGATTAATACCTGAGTACACATCAATATTCCTATATGACAAAGTTCTG	3738
Db	105025	AGCAATTTGGGATGATTAATACCTGAGTACACATCAATATTCCTATATGACAAAGTTCTG	105084
QY	3739	AAACCGATTGTATCTAGTATCTGTTGCTGAGACACCCCATCTTGGATTGGAATCAATCAAGT	3798
Db	105085	AAACCGATTGTATCTAGTATCTGTTGCTGAGACACCCCATCTTGGATTGGAATCAAGT	105144
QY	3799	TATACCTCCAGAGACCTTACCTTCCTAAGCAATTCOCCATTAACCAACGAGATTG	3858
Db	105145	TATACCTCCAGAGACCTTACCTTCCTAAGCAATTCOCCATTAACCAACGAGATTG	105204
QY	3859	ATAGCTATTCAGTATGATTACAGAAATTCATTTGGACGCTATTTTCTCAATTTAAGTTG	3918
Db	105205	ATAGCTATTCAGTATGATTACAGAAATTCATTTGGACGCTATTTTCTCAATTTAAGTTG	105264
QY	3919	GTAATAGTCACATTTGGCTGATAAATAATGACGGGAGGGATACAAATATGAGACATA	3978
Db	105265	GTAATAGTCACATTTGGCTGATAAATAATGAGGAGGGATACAAATATGAGACATA	105324
QY	3979	CAGAATATGGGTAGATTATGATATGACAGTGTATACCCCAATATTTGGCTCACTAAAT	4038
Db	105325	CAGAATATGGGTAGATTATGATATGACAGTGTATACCCCAATATTTGGCTCACTAAAT	105383
QY	4039	CAAAATATTCAAAGCTACAGTATATATTTGGTGGATTGCATATATACCTTGATTTGTTAT	4098
Db	105384	CAAAATATTCAAAGCTACAGTATATATTTGGTGGATTGCATATATACCTTGATTTGTTAT	105443
QY	4099	TATGTACTTGTGCTAGCAATTAATATAGGCTGTTCTTTGGGTGATTTTGTGAAGAAT	4158
Db	105444	TATGTACTTGTGCTAGCAATTAATATAGGCTGTTCTTTGGGTGATTTTGTGAAGAAT	105503
QY	4159	GTTGAGCAAGCACTTCTCCTATTAATGCTAATTTTGTACCTGTTCCAGGAAGCAATCC	4218
Db	105504	GTTGAGCAAGCACTTCTCCTATTAATGCTAATTTTGTACCTGTTCCAGGAAGCAATCC	105563
QY	4219	AATCGGAGTACAGCAGGCGCTCAGAGAAGCAAGGACACCTCACTTGAAATAGCTGGAG	4278
Db	105564	AATCGGAGTACAGCAGGCGCTCAGAGAAGCAAGGACACCTCACTTGAAATAGCTGGAG	105623
QY	4279	GCACAGGTGTATATGTTCAATATGTTATTTTCGATTAAGACATATAAATCCTAATTTACTGG	4338
Db	105624	GCACAGGTGTATATGTTCAATATGTTATTTTCGATTAAGACATATAAATCCTAATTTACTGG	105683
QY	4339	CTACAGATTCAGTATATGATTTACCTGTTTACAGAGTATTCACATTTAAGAGTGGAAATCC	4398
Db	105684	CTACAGATTCAGTATATGATTTACCTGTTTACAGAGTATTCACATTTAAGAGTGGAAATCC	105743
QY	4399	CTGCGTGTAAAGCGCTTCTGATGTTTACCCAGATACATGATGATGCTCTGTTGACAA	4458
Db	105744	CTGCGTGTAAAGCGCTTCTGATGTTTACCCAGATACATGATGATGCTCTGTTGACAA	105803
QY	4459	TGAGTGTGCTAAACCAAGATTTGAGACCTTGAGGACCAAGGATATGCTATATATGCTCTTT	4518
Db	105804	TGAGTGTGCTAAACCAAGATTTGAGACCTTGAGGACCAAGGATATGCTATATATGCTCTTT	105863
QY	4519	GCAATATTCATCCCATGATTTGCTAATTTGGCTGTTTCAAACTTTCACGCTGATGTT	4577
Db	105864	GCAATATTCATCCCATGATTTGGCTAATTTGGCTGTTTCAAACTTTCACGCTGATGTT	105922
QY	4578	GTAACCTGTTTATGAGGATTAATCCCGCATACCAATATATCTGTATAAATTCACGATTTGG	4637
Db	105923	GTAACCTGTTTATGAGGATTAATCCCGCATACCAATATATCTGTATAAATTCACGATTTGG	105982
QY	4638	CCAGTCCAAACATATATTAATCAAGCTCACTGACATCTGACATTAATATGATCAAAATTC	4697
Db	105983	CCAGTCCAAACATATATTAATCAAGCTCACTGACATCTGACATTAATATGATCAAAATTC	106042
QY	4698	CGAGGTGAAGATGCGCAGAGGACTCGGTGAAGCGGTGACAGCATGACGCTTGTTC	4757

Db	106043	GCAGGTAAGAGGACAGAGACTCGGTGAAGCGGGTGAACAGGACATGAAGCCGGTGTTC	106102
QY	4758	CGCGCCTTGTGATATGTCATCCTCAAGATCCATTCACAGCTCCCATCTGAAGCAAC	4817
Db	106103	CGCGCCTTGTGATATGTCATCCTCAAGATCCATTCACAGCTCCCATCTGAAGCAAC	106162
QY	4818	GTCCAGACGCTGTGTCCTCAATCCCAAGTACCCGAACATATACCTTGCCTACTAACAAGA	4877
Db	106163	GTCCAGACGCTGTGTCCTCAATCCCAAGTACCCGAACATATACCTTGCCTACTAACAAGA	106222
QY	4878	CATCGAGGTAAACAACATACATACGCCACATACCTTTTGCGCTCAGAGGACGAGGA	4937
Db	106223	CATCGAGGTAAACAACATACATACGCCACATACCTTTTGCGCTCAGAGGACGAGGA	106282
QY	4938	CTTCGTCAATGGCCGCTGTGGCTCCGGGCAATATGGCCCGGACGCTCCGTGACGGGCT	4997
Db	106293	CTTCGTCAATGGCCGCTGTGGCTCCGGGCAATATGGCCCGGACGCTCCGTGACGGGCT	106342
QY	4998	GGGAGCGCTGAGAGATCTCCAGAGAGAGATGGCGTGGGCGGCGTTCGTCTGGGTGAC	5057
Db	106343	GGGAGCGCTGAGAGATCTCCAGAGAGAGATGGCGTGGGCGGCGTTCGTCTGGGTGAC	106402
QY	5058	GTCTGAGACCCAAACCCAAAGCTGCTCGCTCTGTAAAGACACTCGGAGCCAGAAATGA	5117
Db	106403	GTCTGAGACCCAAACCCAAAGCTGCTCGCTCTGTAAAGACACTCGGAGCCAGAAATGA	106462
QY	5118	TGATGGAGACGCGTAGACATTCGTAATGCTGAGAGTATGATTTGGTGTGCTGCG	5177
Db	106463	TGATGGAGACGCGTAGACATTCGTAATGCTGAGAGTATGATTTGGTGTGCTGCG	106522
QY	5178	TGCTGATGTGCATCTTTGTAAAGGACCTCTTAATACCTCTTCTTAAAGGACTT	5237
Db	106523	TGCTGATGTGCATCTTTGTAAAGGACCTCTTAATACCTCTTCTTAAAGGACTT	106582
QY	5238	AGTACCCCTTGGAGATCTCATCGCTCTCTAAACATACATACATTAACCAATGTGCTGG	5297
Db	106583	AGTACCCCTTGGAGATCTCATCGCTCTCTAAACATACATACATTAACCAATGTGCTGG	106642
QY	5298	TGTGATGGGCTGTGCCAATTTGAAGCACTGATTTGAGGACGCTGATGCTATCTTC	5357
Db	106643	TGTGATGGGCTGTGCCAATTTGAAGCACTGATTTGAGGACGCTGATGCTATCTTC	106702
QY	5358	AATATATTTTGAATGAGCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	5417
Db	106703	AATATATTTTGAATGAGCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	106762
QY	5418	TATACATACCTTTGAGCTGTCTATTTGAAGCAAGCAAGTACATATCTCAGCTGT	5477
Db	106763	TATACATACCTTTGAGCTGTCTATTTGAAGCAAGCAAGTACATATCTCAGCTGT	106822
QY	5478	TCATTTGAAGTTTAAAGTGAATCTGATTAACAAGTACGTAATGTGTAATGAAGGAGC	5537
Db	106823	TCATTTGAAGTTTAAAGTGAATCTGATTAACAAGTACGTAATGTGTAATGAAGGAGC	106882
QY	5538	TGTATACCTCCATATCAATGTCATTTTCATATTAACAATTAAGTATGATCTTTCACCCAA	5597
Db	106883	TGTATACCTCCATATCAATGTCATTTTCATATTAACAATTAAGTATGATCTTTCACCCAA	106944
QY	5598	ATCCCATACAGTTCATACACAGCAAGTACGTAATTAATGTAATTTTCGTTTAC	5657
Db	106944	ATCCCATACAGTTCATACACAGCAAGTACGTAATTAATGTAATTTTCGTTTAC	107002
QY	5658	AGAAAAAAGGAGGCGTTATTAAGATACACGGGGGGGGTGAACAATATACAAATGA	5717
Db	107003	AGAAAAAAGGAGGCGTTATTAAGATACACGGGGGGGGTGAACAATATACAAATGA	107062
QY	5718	ATGAGATCGATCCGTAAGGCGAGCTACGTACACAAAATATGCTAAATCTCCATAC	5777
Db	107063	ATGAGATCGATCCGTAAGGCGAGCTACGTACACAAAATATGCTAAATCTCCATAC	107122
QY	5778	CAACCAAAACAAAGCTTCGACAGGCGCTAAATAGGACGCACTCATCGCTTCGCGGGC	5837
Db	107123	CAACCAAAACAAAGCTTCGACAGGCGCTAAATAGGACGCACTCATCGCTTCGCGGGC	107182

QY	5838	AAGAAACGAAATCAACTGATATACATTGGCAGGGAACCAACAAAGAAAGG-CCATCCAAATCCA	5896
Db	107183	CAGAAACGAAATCAACTGATATACATTGGCAGGGAACCAACAAAGAAAGGCCATCCAAATCCA	107243
QY	5897	ATCCACATCCCAACGCGCATGTGAAGAAGACAGCATGATTCACAGCATATCTTCGCTCTTA	5956
Db	107243	ATCCACATCCCAACGCGCATGTGAAGAAGACAGCATGATTCACAGCATATCTTCGCTCTTA	107303
QY	5957	CAAGTTTGAATACCTTGTACTCTCTCTTCACGGGAAAAAAGAGCATCAGATTATCTGATCT	6016
Db	107303	CAAGTTTGAATACCTTGTACTCTCTCTTCACGGGAAAAAAGAGCATCAGATTATCTGATCT	107363
QY	6017	CGGGCGCGGTTCGATCTTCCTGGGAGATCTGTTGTGGAGTGGCAGAGTGCAGATCGGCT	6076
Db	107363	CGGGCGCGGTTCGATCTTCCTGGGAGATCTGTTGTGGAGTGGCAGAGTGCAGATCGGCT	107423
QY	6077	GCCCCGTTTCTCTTCACGGAACATCGCCACGATGAAGAAGCCAAAGAACAATAATACGCG	6136
Db	107423	GCCCCGTTTCTCTTCACGGAACATCGCCACGATGAAGAAGCCAAAGAACAATAATACGCG	107483
QY	6137	AATGGGAGATCGCCATCTGCATATTAACATTCGATACGGAAGTATTAATACAGAAATGA	6196
Db	107483	AATGGGAGATCGCCATCTGCATATTAACATTCGATACGGAAGTATTAATACAGAAATGA	107543
QY	6197	CATGTAAAGCTGATATTAACGCTGCAGCTT	6227
Db	107543	CATGTAAAGCTGATATTAACGCTGCAGCTT	107573
RESULT 4			
LOCUS	AB021736	5651 bp	DNA
DEFINITION	Oryza sativa (indica cultivar-group) gene for bZIP protein,	linear	PLN 25-DEC-2002
ACCESSION	AB021736		
VERSION	AB021736.1	GI:4115745	
KEYWORDS	bZIP protein.		
SOURCE	Oryza sativa (indica cultivar-group)		
ORGANISM	Oryza sativa (indica cultivar-group)		
REFERENCE	1 Nakase, M., Matsunura, A., Okabe, S., Matsuda, T. and Adachi, T. Gene structure and DNA-binding specificity of the rice seed bZIP protein, RBP		
AUTHORS	2 (bases 1 to 5651)		
REFERENCE	Nakase, M.		
AUTHORS	Direct Submission		
TITLE	Submitted (22-DEC-1998) Masayuki Nakase, Nagoya University, School of Agricultural Sciences, Applied Biological Sciences; Furo-cho, Chikusa-ku, Nagoya, Aichi 464-8601, Japan		
JOURNAL	(E-mail: nakase@agr1.agr.nagoya-u.ac.jp, Tel: 81-52-789-5514, Fax: 81-52-789-4128)		
FEATURES	Location/Qualifiers		
SOURCE	1..5651		
	/organism="Oryza sativa (indica cultivar-group)"		
	/mol_type="genomic DNA"		
	/isolate="IR36"		
	/db_xref="taxon:39946"		
	join(1653..2061,3074..3167,3702..3897,4704..4779,4868..4893,5197..5561)		
	/codon_start=1		
	/product="bZIP protein"		
	/protein_id="GI:4115745"		
	/translation="MERFVSVEELSDPFVWPPPPPOSAAAOOGGGGVAAGGGGVAAGGGGGAANARCSWEYFOKEFLIEAVDSEPVNPSPRREGGIRGAGVVPDYKQDGLSAATATVNDPVRYNAMLKQLEKDLAAVMAKAGTVPPEPAGSLLMDVSHIGAPISGKATPVQNMISPSGSGSSSSVQVAVYVYIKQPTSSSESSSSDDMDKFEBAETGTAPADQSLQRRKOSNBSRKSRSKRAHLNLEDAVQSLQVENSISLRRLLD		
CDS			



VMQYNDANVNRKADYELRAKYVMAEDSYKRYTGNNALFPADSDMSLMPNS  
SPSASIDAAVPIODDPNNYFATNNDJGNNNTMPDIPSSAOEDDEPNALAAKRGIG  
RPSAQIQRVSLHEIQRMCQGPASSGSTS"  
BASE COUNT 1515 a 1163 c 1290 g 1683 t  
ORIGIN

Query Match 74.8%; Score 4658.2; DB 8; Length 5651;  
Best Local Similarity 97.3%; Pred. No. 0;  
Matches 5058; Conservative 0; Mismatches 73; Indels 66; Gaps 29;

```
QY 1 GGFACCCATCTAATACATTAAATACAGAGAAATGATATGCAATATTTATTTT 60
DB 478 GGTACCCATCTAATACATTAAATACAGAGAAATGATATGCAATATTTATTTT 537
QY 61 AAGGAGGCTAATATTTTATCGATTTTAAATACAGGCAATCGGTACTTAGTA 120
DB 538 ATGGAGGCTAATATTTTATCGATTTTAAATACAGGCAATCGGTACTTAGTA 527
QY 121 AAGCTACGTATGACTATCGCTACCGCTACGGTATGATGGAATCTTGATAGCATC 180
DB 598 AAGCTACGTATGACTATCGCTACCGCTACGGTATGATGGAATCTTGATAGCATC 657
QY 181 TGTGTGTGTGTGCACTTGGGTAAGCTTGAATAGCTCCAGCCGTAAGAGAGGTTTTC 240
DB 658 TGTGTGTGTGTGCACTTGGGTAAGCTTGAATAGCTCCAGCCGTAAGAGAGGTTTTC 717
QY 241 GCAAGTTTATAGGATTCGCAAGTTAGACTAGGCAATTCAGTTCAGGATTTGTAG 300
DB 718 GCAAGTTTATAGGATTCGCAAGTTAGACTAGGCAATTCAGTTCAGGATTTGTAG 777
QY 301 TATATGAAAAGAGAGATTCGCAAGTTATATTTATTTATTAAGAGAAATCGACT 360
DB 778 TATATGAAAAGAGAGATTCGCAAGTTATATTTATTTATTAAGAGAAATCGACT 837
QY 361 TGAGGTGTCTAATTCACCAACGAGCTACTCCCGCTTCATATATGATATACATA 420
DB 838 TGAGGTGTCTAATTCACCAACGAGCTACTCCCGCTTCATATATGATATACATA 897
QY 421 TACGTATATACGTATATACATATACATATACATATATG----- 464
DB 898 TACGTATATACGTATATACATATACATATACATATATACATATATATATA 957
QY 465 --GTAATACATATATATATATATATATATATATATATATATATATATATATA 521
DB 958 TATATATATATATATATATATATATATATATATATATATATATATATATATATA 1017
QY 522 GTGGCAATGCTAAAAAGTTTATATATG-AACGATGAAGTCTATCCACT-AACTCC 579
DB 1018 TGGAACATGCTAAAAAGTTTATATATGAAACGATGAAGTCTATCCACTCAAGTCCC 1077
QY 580 TATAGTTTCTGGCACTGTGTGTATGCAATGCAATATATATATATATATATATAT 639
DB 1078 TATAGTTTCTGGCACTGTGTGTATGCAATGCAATATATATATATATATATATAT 1137
QY 640 ATATATTCGTGCGAGCAAAATTAAGACATATATTC-GGTAATACCAATTATACAGATA 698
DB 1138 ATATATTCGTGCGAGCAAAATTAAGACATATATTCAGGATATACCAATTATACAGATA 1197
QY 699 TATCTAATTCACCTGTATATCTAATTAATTCACCTGTATCCCTT-TATGATATATCTG 757
DB 1198 TATCTAATTCACCTGTATATCTAATTAATTCACCTGTATCCCTTGTATGATATATCTG 1257
QY 758 GATAACAATTAATCTACAGAGTATATCTACATATACAGCACTGACACCAACATCCCT 817
DB 1258 GATAACAATTAATCTACAGAGTATATCTACATATACAGCACTGACACCAACATCCCT 1317
QY 818 CACCCAGATGTTACAAAAGCATATTTTGCCTGTATGTTATATATATATATATATAT 877
DB 1318 CACCCAGATGTTACAAAAGCATATTTTGCCTGTATGTTATATATATATATATATAT 1367
QY 878 AAGAGATATGTTAGTATACAAAATTTTCCCAACCCCAACCTCGGAAACAAAACCGGT 937
DB 1368 AAGAGATATGTTAGTATACAAAATTTTCCCAACCCCAACCTCGGAAACAAAACCGGT 1427
```

```
QY 938 AGTATTTGTCTAACACAGAGATCCGATCATTTATTTACACCAACACAAAATTA 997
DB 1428 AGTATTTGTCTAACACAGAGATCCGATCATTTATTTACACCAACACAAAATTA 1487
QY 998 GCAGCAGCAGCCGCTTCCATATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1057
DB 1488 GCAGCAGCAGCCGCTTCCATATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1547
QY 1058 TTCCGCGAGAGTTTCTCCGATCAACCCGATCCGATCCGATCCGATCCGATCCGATCC 1117
DB 1548 TTCCGCGAGAGTTTCTCCGATCAACCCGATCCGATCCGATCCGATCCGATCCGATCC 1607
QY 1118 GTACCCCGGATCTCCGCTGCGAGAGAGGATTCGATTCGCGATGAGAGAGGATTTTC 1177
DB 1608 GTACCCCGGATCTCCGCTGCGAGAGAGGATTCGATTCGCGATGAGAGAGGATTTTC 1667
QY 1178 TCCGTGAGAGATCTCCGATCTCCGATCTCCGATCTCCGATCTCCGATCTCCGATCT 1237
DB 1668 TCCGTGAGAGATCTCCGATCTCCGATCTCCGATCTCCGATCTCCGATCTCCGATCT 1727
QY 1238 GCGGCCAGCAGAGAGGCGGCGGCGGCTTCGAGAGGTGATGATGAGAGG 1297
DB 1728 GCGGCCAGCAGAGAGGCGGCGGCGGCTTCGAGAGGTGATGATGAGAGG 1787
QY 1298 GCGGCGGCGGCGGAGCGATGAAACCGGTGCGGAGTGGTACTTCCAGAGTTT 1357
DB 1788 GCGGCGGCGGCGGAGCGATGAAACCGGTGCGGAGTGGTACTTCCAGAGTTT 1847
QY 1356 CTGAGAGAGGCGGTGCTGATAGCCCGCTCCGAAACCGTAAAGCGAGGCGGAGCGGGA 1417
DB 1848 CTGAGAGAGGCGGTGCTGATAGCCCGCTCCGAAACCGTAAAGCGAGGCGGAGCGGGA 1907
QY 1418 GGGATTCAGGGGCGCAAGAGGCGGTGCGGCTGATGATTAAGACCGCAGCTCTCGCG 1477
DB 1908 GGGATTCAGGGGCGCAAGAGGCGGTGCGGCTGATGATTAAGACCGCAGCTCTCGCG 1967
QY 1478 GCGGCGAGCAGAGCGCGGTGATGACCCGCTGAGATCAAGCGATGCTAAGCAAAAG 1537
DB 1968 GCGGCGAGCAGAGCGCGGTGATGACCCGCTGAGATCAAGCGATGCTAAGCAAAAG 2027
QY 1538 CTGGAGAGAGCTCCCGCGGCTCCGATGAGAGGCTACAGCCATTCGCCCCCTCTA 1597
DB 2028 CTGGAGAGAGCTCCCGCGGCTCCGATGAGAGGCTACAGCCATTCGCCCCCTCTA 2087
QY 1598 GTACTCGAGAGCTTATGAGATCGGCAATGCTTACGTAAGTTCGATGGAATGTTTATG 1657
DB 2088 GTACTCGAGAGCTTATGAGATCGGCAATGCTTACGTAAGTTCGATGGAATGTTTATG 2147
QY 1658 GTATTGATCGGGCATTTCTTATAGCAATGGGCTGCANTGCTTGCATAGGCTCTGT 1717
DB 2148 GTATTGATCGGGCATTTCTTATAGCAATGGGCTGCANTGCTTGCATAGGCTCTGT 2207
QY 1718 TTAGTGTGCTGTGTTGCTGATCAGTATGATGAGGTTGTGCTATGTACAAACGAA 1777
DB 2208 TTAGTGTGCTGTGTTGCTGATCAGTATGATGAGGTTGTGCTATGTACAAACGAA 2261
QY 1778 GCTTCATAGACCTCGGATATGATGATGATGATGATGATGATGATGATGATGATG 1835
DB 2266 GCTTCATAGACCTCGGATATGATGATGATGATGATGATGATGATGATGATGATG 2321
QY 1836 TAAATCAGTATATCTTAACTAGCAGACAGTGTGTTGTTGCAATTCGAATATATTT 1895
DB 2322 TAAATCAGTATATCTTAACTAGCAGACAGTGTGTTGTTGCAATTCGAATATATTT 2381
QY 1896 TGTATGAGGCTGCTTAACTGT-CATGCTCTTTTATGATGCAATATGATCTTATGTC 1954
DB 2382 TGTATGAGGCTGCTTAACTGTATGATCTCTTTTATGATGCAATATGATCTTATGTC 2441
QY 1954 AAGTATTT-TGATGCGCGGATCCAGAACAAAAGTTGGGGGATTCACATACCGAGTA 2013
DB 2442 AAGTATTTATGATGCGCGGATCCAGAACAAAAGTTGGGGGATTCACATACCGAGTA 2501
```





Db 4656 GTTGAAGCAAGCACTCTTCGATATAATGCTAGTTTGTGACCTGTTCAGAGAACATCC 4715  
 Oy 4219 AATGGGAGTCAGCAGCGCTCAGAGAGCAAGAGGACCTCAGCTGAAAGACCTGAG 4278  
 Db 4716 AATGGGAGTCAGCAGCGCTCAGAGAGCAAGAGGACCTCAGCTGAAAGACCTGAG 4775  
 Oy 4279 GCACAGGTGTAGTACATGATAGTATTTTTCATAGAGATATAATCCCTAAATTTAGTG 4338  
 Db 4776 GCACAGGTGTAGTACATGATAGTATTTTTCATAGAGATATAATCCCTAAATTTAGTG 4835  
 Oy 4339 CTACGACTGATGATGATGATTTACTCTGTTCAGAGTTCGCAATTAAGAGTCGAAATC 4398  
 Db 4836 CTACGACTGATGATGATGATTTACTCTGTTCAGAGTTCGCAATTAAGAGTCGAAATC 4895  
 Oy 4399 CTGCGTTTAAAGCGGTGCTGATGATTAACAGAGTACATGATGATGCTGCTGACAA 4458  
 Db 4896 CTGCGTTTAAAGCGGTGCTGATGATTAACAGAGTACATGATGATGCTGCTGACAA 4555  
 Oy 4459 TAGAGTCTAAAGCAAGATGTGAGACCTTGAAGCAAGGTATGCTATATATGCTTTT 4518  
 Db 4956 TAGAGTCTAAAGCAAGATGTGAGACCTTGAAGCAAGGTATGCTATATATGCTTTT 5015  
 Oy 4519 GCAATATGCAATCCCATGAGAT--GCTACTTGGCTGTTCCTAACTTTCAAGCTACTGT 4577  
 Db 5016 GCAATATGCAATCCCATGAGATGCTACTTGGCTGTTCCTAACTTTCAAGCTACTGT 5074  
 Oy 4578 GTACCTGTATTTAGAGATATATCCCGCTACCATATATCTATATAATCCATTTGG 4637  
 Db 5075 GTACCTGTATTTAGAGATATATCCCGCTACCATATATCTATATAATCCATTTGG 5134  
 Oy 4638 CCAATGCCAATGATATTAATTAACAGCTCAATCTGAACATGATAATGATCAAAATTC 4697  
 Db 5135 CCAATGCCAATGATATTAATTAACAGCTCAATCTGAACATGATAATGATCAAAATTC 5194  
 Oy 4698 GGAAGTGAAGTGCAGAGGACGCTGAGAGCGGGTACAGGATGAAGGCTGTTTCC 4757  
 Db 5195 GGAAGTGAAGTGCAGAGGACGCTGAGAGCGGGTACAGGATGAAGGCTGTTTCC 5254  
 Oy 4758 CGCGCTTCTGATATGTCATCCCTCAGATGCCATTCACAGCTCCCATCTGAAGCAAC 4817  
 Db 5255 CGCGCTTCTGATATGTCATCCCTCAGATGCCATTCACAGCTCCCATCTGAAGCAAC 5314  
 Oy 4818 GTACAGCGCTGCTGTTCCTCAATGCAAGTACCGCAATTAATTTGGCTACACAGCA 4877  
 Db 5315 GTACAGCGCTGCTGTTCCTCAATGCAAGTACCGCAATTAATTTGGCTACACAGCA 5374  
 Oy 4878 CATCGAGGTAAACACACTACATGACCGCAATCTCTTCGCTCAGAGAGAGAGA 4937  
 Db 5375 CATCGAGGTAAACACACTACATGACCGCAATCTCTTCGCTCAGAGAGAGAGA 5434  
 Oy 4938 CTTCCTCAATGGCGCTGTGGCTCCGCGCAAGATTGGCCGCGCAGCTCGCTCAGCGGCT 4997  
 Db 5435 CTTCCTCAATGGCGCTGTGGCTCCGCGCAAGATTGGCCGCGCAGCTCGCTCAGCGGCT 5494  
 Oy 4998 GCGAGAGCTTGAGCATTCACAGAGAGATGGGCGGCGGCTTCGTCGGGTCGAC 5057  
 Db 5495 GCGAGAGCTTGAGCATTCACAGAGAGATGGGCGGCGGCTTCGTCGGGTCGAC 5554  
 Oy 5058 GTCTGAGACCGAAACCCAGAGCTGCTTGGTTCTGAAGAACACTCGAGCAGAAATGA 5117  
 Db 5555 GTCTGAGACCGAAACCCAGAGCTGCTTGGTTCTGAAGAACACTCGAGCAGAAATGA 5614  
 Oy 5118 TGATTGGAAGAGGCTAGACATTTGCTTAATGCTGAGG 5154  
 Db 5615 TGATTGGAAGAGGCTAGACATTTGCTTAATGCTGAGG 5651

RESULT 5  
 AY072929 1065 bp mRNA linear PLN 17-FEB-2002  
 LOCUS  
 DEFINITION Oryza sativa bZIP protein (bZIP) mRNA, partial cds.  
 ACCESSION AY072929 GI:18698990  
 VERSION AY072929.1

KEYWORDS  
 SOURCE  
 ORGANISM  
 Oryza sativa  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Euhartioideae; Oryzoideae; Oryza.  
 REFERENCE  
 1 (bases 1 to 1065)  
 Qianhong Y., Rihe P., Aisheng X., Xian L. and Huijin F.  
 AUTHORS  
 DIRECT SUBMISSION  
 JOURNAL  
 Submitted (11-JAN-2002) Bio-tech, Shanghai Academy of Agriculture  
 Science, Beid Road 2901, Shanghai 201106, China  
 FEATURES  
 source  
 1..1065  
 /organism="Oryza sativa"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4530"  
 /note="Isolated by yeast one-hybrid system using conserved  
 C-box encoded by Genbank Accession Number AF405700 as  
 bait"  
 <1..1065  
 /gene="bZIP"  
 <1..723  
 /gene="bZIP"  
 /note="bZIP; transcription factor; contains C-box binding  
 domain; similar to plant genes regulated by diverse  
 environmental, physiological, and environmental cues"  
 /codon\_start=1  
 /product="bZIP protein"  
 /protein\_id="AA177201.1"  
 /db\_xref="GI:18698991"  
 /translation="GGGSGSOLVQNDVIVKOPTSSSRSSRESDDDMEGAEATTGAR  
 PADORLRKQSNRESARSRSRKALNELAEVQSLRVNSLRLRLDVOKND  
 AAVDNVILKADVETLRTVKMAEDSVRYVGNMLFPAADMSDFSPNDSSEAT  
 DAAPVDDPNVYFATNNDIGNNVNPDIPSAQEEDFPNGALAGKIGRPDLQR  
 VALLEILQRMGGSPASSGSTS"  
 BASE COUNT  
 290 a 245 c 279 g 251 t  
 ORIGIN  
 Query Match 10.6%; Score 661; DB 8; Length 1065;  
 Best Local Similarity 97.8%; Pred. No. 4.4e-130;  
 Matches 670; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 5180 GTCATTGTCATTCTTTGTAAGGACACCTCTAGTACCTCTTCTTCTAAGGACTTAG 5239  
 |||||  
 DB 839 GTCATTGTCATTCTTTGTAAGGACACCTCTAGTACCTCTTCTTAAAGGACTTAG 898  
 |||||  
 QY 5240 TACCCCTGTGTGATCAGTCTGCTAACTAATACATTAACCAATGTTCAATGTTG 5299  
 |||||  
 DB 899 TACCCCTGTGTGATCAGTCTGCTAACTAATACATTAACCAATGTTCAATGTTG 958  
 |||||  
 QY 5300 TGATGGCGTCCCTCAATTTGTAAGACGATTTCAAGGACGCGTGTGATGATCA 5359  
 |||||  
 DB 959 TGATGGCGTCCCTCAATTTGTAAGACGATTTCAAGGACGCGTGTGATGATCA 1018  
 |||||  
 QY 5360 TAATATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5384  
 |||||  
 DB 1019 TAATATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1043  
 |||||

RESULT 6  
 RICEZIPPA 1696 bp mRNA linear PLN 23-MAR-2002  
 LOCUS Oryza sativa (japonica cultivar-group) mRNA for bZIP protein,  
 complete cds.  
 ACCESSION D78609.1 GI:1122224  
 VERSION D78609  
 KEYWORDS bZIP protein.  
 SOURCE Oryza sativa (japonica cultivar-group)  
 ORGANISM Oryza sativa (japonica cultivar-group); Embryophyta; Tracheophyta;  
 Eukaryota; Viridiplantae; Streptophyta; Liliopsida; Poales; Poaceae;  
 Eriarthroideae; Oryzaceae; Oryza.

REFERENCE 1  
 AUTHORS Nakase, M., Aoki, N., Matsuda, T. and Adachi, T.  
 TITLE Characterization of a novel rice bZIP protein which binds to the  
 alpha-globulin promoter  
 JOURNAL Plant Mol. Biol. 33 (3), 513-522 (1997)  
 MEDLINE 97201485  
 PUBMED 9049271

REFERENCE 2 (bases 1 to 1696)  
 AUTHORS Nakase, M., Adachi, T., Matsuda, A. and Matsuda, T.  
 TITLE Characterization of a novel rice bZIP protein which binds to the 5'  
 region of the alpha-globulin gene  
 JOURNAL Unpublished  
 REFERENCE 3 (bases 1 to 1696)  
 AUTHORS Nakase, M.  
 TITLE Submitted (07-DEC-1995) Masayuki Nakase, Nagoya University, School  
 of Agricultural Sciences, Applied Biological Sciences, Furo-cho,  
 Chikusa-ku, Nagoya, Aichi 464-01, Japan  
 (E-mail:145231a@nuc.cc.nagoya-u.ac.jp, Tel:052-789-4131,  
 Fax:052-789-4120)

FEATURES  
 source location/Qualifiers  
 1..1696  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /mol\_type="mRNA"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:39947"  
 /tissue\_type="endosperm"  
 /dev\_stage="endosperm, immature"  
 146..1423  
 /codon\_start=1  
 /product="bZIP protein"  
 /protein\_id="BA11431.1"  
 /db\_xref="GI:1783305"  
 /translation="MERVSVVEISIDPPWPPPPPOSAAAQOOGGSGVAGSGGGA  
 GGGGNNMRCPSEMYFQKFLFEAVLDSPYPPSRRAAGIRAGGVVVDYKPO  
 LSAHAATTSAYVDPEYNNMKOKIEDLAAMWASTVPPPEPAGSLINA  
 DYSHGAPISGKATPPVNMISGSGSGSQLYQVVDVIVAGVSSSSRQSDDDW  
 BEBATTGTAPDQRLQRRKOSNRSRSRKAHLNELAQVSQLRVSSSLR  
 RLADNMKYNDAAVDNRVLKADVELTLRAKYMAEDSVKVTGMNALLFPASDMSLSM  
 PPNSSPSTDAVPIODPNNTATNDIGNNYMDIPSSAQEDEFVNGALAA  
 KIGRTASLRVAVLEHLQKRWGGPAGSGST"

BASE COUNT 362 a 461 c 532 g 341 t

## ORIGIN

Query Match 9.0%; Score 563.4; DB 8; Length 1696;  
 Best Local Similarity 99.8%; Pred. No. 2.9e-109;  
 Matches 564; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4700 AGGTGAAGATGCGACAGGACTCGGTGAAGGCGGTGACAGGATTAACGCTTGTCCG 4759  
 |||||  
 DB 1059 AGGTGAAGATGCGACAGGACTCGGTGAAGGCGGTGACAGGATTAACGCTTGTCCG 1118  
 |||||  
 QY 4760 CCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4819  
 |||||  
 DB 1119 CCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1178  
 |||||  
 QY 4820 CAGAGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4879  
 |||||  
 DB 1179 CAGAGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1238  
 |||||  
 QY 4880 TCGAGGTGAACAACATTAACATTAACATTAACATTAACATTAACATTAACATTA 4939  
 |||||  
 DB 1239 TCGAGGTGAACAACATTAACATTAACATTAACATTAACATTAACATTAACATTA 1298  
 |||||  
 QY 4940 TCGATGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4999  
 |||||  
 DB 1239 TCGATGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1358  
 |||||  
 QY 5000 CGAGCTGAGCATCTCCAGAAAGAGATGTCGCTGCTGCTGCTGCTGCTGCTGCTG 5059  
 |||||  
 DB 1359 CGAGCTGAGCATCTCCAGAAAGAGATGTCGCTGCTGCTGCTGCTGCTGCTGCTG 1418  
 |||||  
 QY 5060 CCTGAGACCGAACCAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5119  
 |||||  
 DB 1419 CCTGAGACCGAACCAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1478  
 |||||  
 QY 5120 ATTGACAGGCGTAGACATTCCTAATGCTGAGGTTGATGATGATGATGATGATG 5179  
 |||||  
 DB 1479 ATTGACAGGCGTAGACATTCCTAATGCTGAGGTTGATGATGATGATGATGATGATG 1538  
 |||||  
 QY 5180 GTCATTGTCATTCTTTGTAAGGACACCTCTAGTACCTCTTCTTCTAAGGACTTAG 5239  
 |||||  
 DB 1539 GTCATTGTCATTCTTTGTAAGGACACCTCTAGTACCTCTTCTTCTAAGGACTTAG 1598  
 |||||

QY 5240 TACCCCTGTGTGATCAGTCTGCTAACTAATACATTAACCAATGTTCAATGTTG 5264  
 |||||  
 DB 1599 TACCCCTGTGTGATCAGTCTGCTAACTAATACATTAACCAATGTTCAATGTTG 1623  
 |||||

RESULT 7  
 HVB121 4027 bp DNA linear PLN 01-MAR-1997  
 LOCUS H. vulgare b1z-1 gene.  
 DEFINITION X80068  
 ACCESSION X80068  
 VERSION X80068.1 GI:1869927  
 KEYWORDS b1z-1 gene; b1z-1 protein.  
 SOURCE Hordeum vulgare subsp. vulgare  
 ORGANISM Hordeum vulgare subsp. vulgare  
 Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Pooideae; Triticeae; Hordeum.

REFERENCE 1  
 AUTHORS Vicente-Carvajosa, J., Lama, B., Gonzalez, H., Orate, L. and  
 Cardonero, P.  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 4027)  
 AUTHORS Vicente-Carvajosa, J.  
 TITLE Direct Submission  
 JOURNAL Submitted (07-JUL-1994) J Vicente-Carvajosa, ETS Ingenieros  
 Agronomos, Universidad Politecnica Madrid, Dept de Biologia,  
 Diagonal Universitaria s/n, 28040 Madrid, Spain

FEATURES  
 source location/Qualifiers  
 1..4027  
 /organism="Hordeum vulgare subsp. vulgare"  
 /mol\_type="genomic DNA"

```

/cultivar="Sunbar"
/db_xref="taxon:112509"
/chromosome="5H"
join(34, .360,1602, .1697,2147, .2343,3017, .3092,3181, .3306,
3509, .3862)
/gene="b12-1"
join(34, .360,1602, .1697,2147, .2343,3017, .3092,3181, .3306,
3509, .3862)
/gene="b12-1"
/codon_start=1
/product="b12-1 protein"
/protein_id="CA56374.1"
/db_xref="GI:1869928"
/db_xref="SPTREMBL:P93667"
/translation="MERVSVSEIIPDEFGOPSPROGRHREPPEGANMPCSEWTFQKF
LEAVLDSPADSPKSGASGGOACRPRGAGTGTGPAVDPEYNAAMKQKLEKL
AAVAMRAVSAMPPEPRFASBSCPNAGOHITGINIGSNVYPILOKLAGASGVSGP
HIVONRADLVKQASSSRESEDDDEGDELITNGVPTDRLRRKQSRRESARVS
RSRAAHLELEAQSOLKEVENSILRLADVOKYNGAIVNRYLADEVLELRAVK
MAEDSVKRYGMALFPASDMSLSMPFGSPSEADAFDPDLSAFTSESEAGN
NGYMEASSAODDFINETMDTSKMGPRDSLHRYASLEHLQKMGSGPSSSIS"
34, .360
/gene="b12-1"
/number=1
361, .1601
/gene="b12-1"
/number=1
1602, .1697
/gene="b12-1"
/number=2
1698, .2146
/gene="b12-1"
/number=2
2147, .2343
/gene="b12-1"
/number=3
2344, .3016
/gene="b12-1"
/number=3
3017, .3092
/gene="b12-1"
/number=4
3093, .3180
/gene="b12-1"
/number=4
3181, .3306
/gene="b12-1"
/number=5
3307, .3508
/gene="b12-1"
/number=5
3509, .3862
/gene="b12-1"
/number=6
3863, .4199

```

Query Match 6 58: Score 402; DB 8; Length 4027;  
 Best Local Similarity 65.88; Pred. No. 8.3e-75;  
 Matches 634; Conservative 0; Mismatches 320; Indels 10; Gaps 3;

Db 4139 GGGTGAATTTTATGAAGATTTGAGCAAGCACTCTCGATATATAGTATGTTGTTGA 4198  
 2947 GGGTCTCTCTATAGAGCTATTTAGCGTCCACTCTTGATATATAGTATGTTGTTGA 3006  
 QY 4199 CCTGTTCAGAGCAATCCATATCGGAGTCAAGCCAGGCGCTCAAGACAGAAAAGCGAG 4258  
 Db 3007 CTGGTTCAGAGCAATCCATATCGGAGTCAAGCCAGGCGCTCAAGACAGAAAAGCGAG 3066  
 QY 4259 CTCACCTGAATGAAGTGAAGCAAGGAGTGTAGTATGATATGTTTGTGATTAACAG 4318  
 Db 3067 CTCACCTGAATGAAGTGAAGCAAGGAGTGTAGTATGATATGTTTGTGATTAACAG 3126

```

QY 4319 ATAAATCCTAAATTAAGTCTACTGACTCACTGATATGATATTAAGTTTACAGATACG 4378
    3127 AGTGATATTCGAAAAAGGCGCTAATGAGTTTGTGTGATGATGAGGCTGTTACAGATAC 3186
    QY 4379 CAATTAAGAGTCGAGTACTCTGCTGTGTAGGCGCTGTGATGTTAAACGAAAGTAC 4438
    Db 3187 CAGTTAAGAGTGAAGAACTCTGCTGTGTAGGCGCGCTGTGATGTTAAACGAAAGTAC 3246
    QY 4439 AATAGTCGCTGTTGACATATGAGTCTAAAGCGAGATGTTGAGACCTGAGCGAAG 4498
    Db 3247 AATGTCCTCTCTTACATATGAGTCTTAAAGCGAGATGTTGAGACCTGAGCGAAG 3306
    QY 4499 GTATGCTATATATGCTTTTGAATATGATATGATATGATATGATATGATATGAT 4558
    Db 3307 GTATGCTATATATGCTTTTGAATATGATATGATATGATATGATATGATATGAT 3365
    QY 4559 AACTTTAAGTACTGTGTACCTGTATTTATGAAAGATTAATCCGCTTACATATAC 4618
    Db 3366 TAGCTTTGTTCCAAATGATGATGATGATGATGATGATGATGATGATGATGAT 3425
    QY 4619 TCTAATATCAGATTTGCGCAGTCCAAATGATATTAATGATATGATATGATATG 4678
    Db 3426 GAATAGCTGCTAGTATGATATGATATGATATGATATGATATGATATGATATGAT 3485
    QY 4679 TGAATGTATCAAAATTTGCGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 4738
    Db 3486 TACTTCCACAGAAATTTGCGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 3545
    QY 4739 GCATGAGCGCTGTGTTCCGCGCTGCTGATATGATATGATATGATATGATATGAT 4798
    Db 3546 GCATGAGCGCTGTGTTCCGCGCTGCTGATATGATATGATATGATATGATATGAT 3605
    QY 4799 GCTCCCATCTGAGCAAGTCAAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4858
    Db 3606 GCTCCCATCTGAGCAAGTCAAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3659
    QY 4859 ACTTCGCTACTAACAAGCAATCGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 4918
    Db 3660 ACTTCGCTACTAACAAGCAATCGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3716
    QY 4919 CGGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4978
    Db 3716 CGGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3776
    QY 4979 CAGCTCGCTGCAAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5038
    Db 3776 CAGCTCGCTGCAAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3836
    QY 5039 CGGCTCGCTGCAAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5098
    Db 3837 CAGCTCGCTGCAAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3896
    QY 5099 CACT 5102
    Db 3897 AACT 3900

```

RESULT 8  
 ZM035063 1453 bp mRNA PLN 25-NOV-1996  
 LOCUS ZM035063  
 DEFINITION Zea mays opaque-2 heterodimerizing protein 1b (Ohp1b) mRNA,  
 complete cds.  
 ACCESSION U35063  
 VERSION U35063.1 GI:1144535  
 KEYWORDS  
 ORGANISM Zea mays  
 Zea mays  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoidae; Andropogoneae; Zea.  
 1 (bases 1 to 1493)  
 REFERENCE Pysht.L.D. and Schmidt,R.J.  
 AUTHORS Characterization of the maize Ohp1 gene: evidence of gene copy  
 TITLE



```

QY 4927 GAGACAGAGACTTCGTCATGCGCTGCGCGCAAGATTGCGCGGCGACCTCG 4986
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1212 GAGACAGAGACTTCGTCATGCGCGCTGCGCGCAAGATTGCGCGGCGACCTCG 1271

QY 4987 CTGACAGCGGCTGCGCGAGCTGAGCATCTCCAGAAAGATGTGCGGTGCGCGCTTCG 5046
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1272 CTGACAGCGGCTGCGCGAGCTGAGCATCTCCAGAAAGATGTGCGGTGCGCGCTTCG 1331

QY 5047 TCTGGTCGAGCTCTGAGACCGAAGCA 5076
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1332 TCGGGTCGAGCTCTGAGACCGAAGCA 1361

RESULT 10
MZEBZIP 1689 bp mRNA linear PLN 16-OCT-2001
LOCUS Zea mays opaque2 heterodimerizing protein 2 mRNA, complete cds.
DEFINITION L06478
ACCESSION L06478 GI:168427
VERSION L06478.1
KEYWORDS Zea mays
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
Clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 1689)
Pysh,L.D., Aukerman,M.J. and Schmidt,R.J.
OHPI: a maize basic domain/leucine zipper protein that interacts
with opaque2
JOURNAL Plant Cell 5 (2), 227-236 (1993)
MEDLINE 93200807
PUBMED 8453304

FEATURES
source
    location/Qualifiers
    1..1689
    /organism="Zea mays"
    /mol_type="mRNA"
    /strain="R802"
    /db_xref="taxon:4577"
    /tissue_type="endosperm"
    /tissue_lib="R802 02-676"
    148..1380
    /codon_start=1
    /product="opaque2 heterodimerizing protein 2"
    /protein_id="AA33439.1"
    /db_xref="GI:168428"
    /translation="MERVSMEEIPNPYNAPPHPOPAAGAVAAAGVGAGADAGAN
    NRCPSEWYEFKLEAVLIDSPGPAVGRRSGAGVLAESPGLGAAPASVSSVVD
    PVENYNAIKOKLEKDLAIAWWRASGAAPDLATASIPSYGVHAPAPKPYGTRES
    LYONMLAGAPVGGSGPHIVQADIPIVQOTSSSREPOSDDDMDGDEFYNGNPGPO
    RQQRKRSNRESARSRSRKAHNEBAVQALTRKNSLLRLADYNORNEAVD
    NRVLKADVEFLTKVYMAEDSVKRVTMNLVPAVDSLSMPNPSFDSASDSV
    PVQDLNLSYFANSEISGNGMPDIASVQDDNFVNGYQAAGKMRSTLSLORVASL
    EHLQKRMCGGPASSGSTS"

BASE COUNT 412 a 446 c 459 g 372 t
ORIGIN

Query Match 3.98; Score 244; DB 8; Length 1689;
Best Local Similarity 68.88; Pred. No. 3,4e-41;
Matches 351; Conservative 0; Mismatches 155; Indels 4; Gaps 1;

```

```

QY 4867 ACTAACAAGACATGAGAGTAACACAACTACATCCCGACATACCTCTTCGCTCAG 4926
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1183 AACCAAGAGAAATGAGAGGACACAAAGCTTAAGCCAGATATAGCTCTCTCGTTCA 1242

QY 4927 GAGACAGAGACTTCGTCATGCGCTGCGCGCAAGATTGCGCGGCGACCTTCG 4986
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1243 GAGACAGAGACTTCGTCATGCGCTGCGCGCAAGATTGCGCGGCGACCTTCG 1302

QY 4987 CTGACAGCGGCTGCGCGAGCTGAGCATCTCCAGAAAGATGTGCGGTGCGCGCTTCG 5046
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1303 CTGACAGCGGCTGCGCGAGCTGAGCATCTCCAGAAAGATGTGCGGTGCGCGCTTCG 1362

QY 5047 TCTGGTCGAGCTCTGAGACCGAAGCA 5076
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1363 TCGGGTCGAGCTCTGAGACCGAAGCA 1361

QY 5103 GCGACAGAGAAATGATGATGAGACGCGTACGATTCGTCATGCGGTGATGAT 5162
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1423 TAAATTACAAATTCACAACTCTTGTACGACATGATTAATCCTGCGANGCGCATTAG 1482

QY 5163 TGTGTGTCGTCGTCGTCGTCATTCGATT 5192
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1483 TCGCTTATGACTTGTGCGTCATTCGATT 1512

RESULT 11
CNS08C70 108982 bp DNA linear HTG 12-MAR-2003
LOCUS Oryza sativa (japonica cultivar-group) chromosome 12 clone
DEFINITION OJ1298_D03, *** SEQUENCING IN PROGRESS ***.
ACCESSION AL731741.4 GI:28950589
VERSION AL731741.4
KEYWORDS HTG; HTGS_PHASE2; HTGS_ACTIVEFIN.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 108982)
Chouine,N., Orjeda,G., Catolico,L., Demange,N., Wincker,P.,
Seuens,B., Pelletier,E., Scarpell,C., Salanoubat,M.,
Weissenbach,J. and Queller,F.
Oryza sativa chromosome 12 sequencing
Unpublished
2 (bases 1 to 108982)
Genoscope.
Direct Submission
Submitted (10-MAR-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
On Mar 13, 2003 this sequence version replaced gi:28460559.
Center: genoscope / Centre National de Sequencage
Contact code: GS
Web site: http://www.genoscope.cns.fr/
Contact: SeqRef@genoscope.cns.fr
-----
The following sequence is oriented from the T7 to the SP6 end.
IMPORTANT: This sequence is unfinished and does not necessarily
represent the correct sequence. Work on the sequence is in progress
and the release of this data is based on the understanding that the
sequence may change as work continue. The sequence may be
contaminated with foreign sequence from E.coli, yeast, vector,
phage, etc. . . even if efforts are made to eliminate these
contaminating sequences.
The nucleotide sequence of this BAC clone was generated by
combining Monsanto, Syngenta and Genoscope sequencing data.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
Location/Qualifiers
1..108982
/organism="Oryza sativa (japonica cultivar-group)"
FEATURES
source

```



AUTHORS  
TITLE  
JOURNAL

## COMMENT

Sasaki, T., Matsumoto, T. and Yamamoto, K.  
Direct Submission  
Submitted (06-DEC-2000) Takuji Sasaki, National Institute of  
Agricultural Sciences, Rice Genome Research Program, Kannondai  
2-1-2, Tsukuba, Ibaraki 305-8602, Japan  
(E-mail: tsasaki@nias.affrc.go.jp, URL: http://rtp.dna.affrc.go.jp/  
Tel: 81-298-38-7441, Fax: 81-298-38-7468)  
On Mar 28, 2001 this sequence version replaced gi:11602827.  
Genes were predicted from the integrated results of the following:  
GENSCAN1.0, BLASTN2.0, BLASTX2.0 as well as SplicePredictor  
(October 1998 version). The genomic sequence was searched against  
NCBI NonRedundant Protein database, or  
(ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at  
RGP. Protein homologues of the coding regions were searched against  
NCBI NonRedundant Protein database with BLASTP2.0. ESTs represent  
the identified cDNA sequences using BLASTN 2.0 with the  
corresponding DBJ accession no. and RGP clone ID.  
A gene with identity or significant homology to a protein is  
classified based on the protein name to indicate the homology level  
such as same name, 'putative-' and '-like protein'. A gene without  
significant homology to any protein but with EST homology (covering  
almost the entire length of partial sequence) is classified as an  
'unknown' protein. A gene predicted with a gene prediction program  
is classified as a 'hypothetical' protein.  
The orientation of the sequence is from SP6 to T7 of the PAC clone.  
This sequence of P0498A12 clone has an overlap with OSJNBa004B13  
(DBJ:AP003018) clone at the position 1 to 82,983 of 5' end. The  
sequence of this clone starts at the position 59,286 of  
OSJNBa004B13. This sequence of P0498A12 clone has an overlap with  
P0511C01 (DBJ: AP002070) clone at the position 158,595 to 159,749  
of 3' end. The sequence of this clone ends at the position 1,155 of  
P0511C01. Detailed information on overlap and assembly quality  
together with annotation of this entry is available at  
http://rtp.dna.affrc.go.jp/genomeseq.html.  
Location/Qualifiers

FEATURES  
source

1. 159749  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="genomic DNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/chromosome="1"  
/clone="P0498A12"  
join(1..65,132..333)  
/gene="P0498A12.1"  
join(1..65,132..333)  
/gene="P0498A12.1"  
/note="hypothetical protein"  
/codon\_start=1  
/protein\_id="BAB39973.1"  
/db\_xref="GI:13486739"  
/translation="MEMGYSSESRKGPSPKDISVVTATARRRCGGDEHTAAR  
ASYIHGDHHLGGIGKVVDSAGAGVRYTVERNDIFYCEQENTRGK"  
complement(join(1424..2215,3205..3465))  
/gene="P0498A12.2"  
complement(join(1424..2215,3205..3465))  
/gene="P0498A12.2"  
/note="hypothetical protein  
similar to Arabidopsis thaliana chromosome 3, MKP6.17"  
/codon\_start=1  
/protein\_id="BAB39974.1"  
/db\_xref="GI:13486740"  
/translation="MGAGKSSGRSPFSGGGGGVGRPRGMLPLAQLVLEKRG  
ASRPPTAALANALYLRPGSLDALPISNVAFPNPHLIHYCDLRFPLSAFHL  
SEHFFPMSSILMKGIQLETSMSVEPAAVALLAGSGGITLLISGLISGNEBA  
YVDYATGPGSGVLEGMKVTANMSDVFYFHGYITPAKMAAEELLILQATPESTSI  
CHLGGITAGLATIMLKRSFGPDPISLISIGTGAVRMPVGVQVLFSGRPGYTSS  
RGVRGSAARENGRGIWCSACTYDNSFDIDCEMSARDEAHAFSHROHQAQNGNE  
PSVEIRRRRLERFSR"  
join(7764..7884,8125..8279,8609..8791,9005..9357,  
9534..9702)  
/gene="P0498A12.3"  
join(7764..7884,8125..8279,8609..8791,9005..9357,  
9534..9702)  
CDS

/gene="P0498A12.3"  
/note="contains ESTs A0058421(E51134), A0031151(E61006)"  
/codon\_start=1  
/product="putative cinnamoyl-CoA reductase"  
/protein\_id="BAB39975.1"  
/db\_xref="GI:13486741"  
/translation="MYGVDEQWCVYGAAGGFTGSLVYKELLHRGYAAVAARVDPDEGRK  
NAHLHLERAKRRLSLHRADVLDCNSLRAFNICDGFHVASPVSDDEPLPLAIEGT  
KNYINAAAMGKRVYFTSSGAAHNNPNSRSDQFLDEGMSDLEPCOTOMNYCYAK  
TVAEKATTEASKRGVOLLVYPAVYAVYGMLOPTLNASVYVAATYMGRTKSAVPAVA  
AYVDYDVARAALYEPDPDRGRTICISVYLRBSEFRLRLREFPOYPTTTCEDNS  
KPMVKPYQSVORLELGMQFTPKESLYRYISQDGHLPALSPRSAL"  
join(12258..12381,13185..13339,13451..13633,14762..15114,  
15350..15521)  
/gene="P0498A12.4"  
join(12258..12381,13185..13339,13451..13633,14762..15114,  
15350..15521)  
/gene="P0498A12.4"  
/note="contains EST A0095269(E61006)"  
/codon\_start=1  
/product="putative cinnamoyl-CoA reductase"  
/protein\_id="BAB39976.1"  
/db\_xref="GI:13486742"  
/translation="MYGRSPQWCVYGAAGGFTGSLVYKELLHRGYAAVAARVDPDEGRK  
NAHLHLERAKRRLSLHRADVLDCNSLRAFNICDGFHVASPVSDDEPLPLAIEGT  
KNYINAAAMGKRVYFTSSGAAHNNPNSRSDQFLDEGMSDLEPCOTOMNYCYAK  
KMAEPTRAEAEASKRGVNLVYPAVYAVYGMLOPTLNASVYVAATYMGRTKSAVPAVA  
AAVYDVARAALYEPDPDRGRTICISVYLRBSEFRLRLREFPOYPTTTCEDNS  
SKPMVKPYQSVORLELGMQFTPKESLYRYISQDGHLPALSPRSAL"  
complement(join(22213..22217,22550..22759,22937..2315))  
/gene="P0498A12.5"  
complement(join(22213..22217,22550..22759,22937..2315))  
/gene="P0498A12.5"  
/note="hypothetical protein"  
/codon\_start=1  
/protein\_id="BAB39977.1"  
/db\_xref="GI:13486743"  
/translation="MYGGRMAKGVGSIDAVYSAGKGRPRMSAAVNGYFISVGRGR  
RGAASAAEASSRGVYITPRYMPGPGPTGARPRRERMAEBCASITARTTHERAVPCQ  
PCTRTGPGTQAQHASGREGTGMHPLTCLVGLDWPVTPQSPSVGRRRHPGPAVGP  
GDRPDPSPRPPCLVASGLMAITPRMGCGVATRV"  
27792..28895  
/gene="P0498A12.6"  
27792..28895  
/gene="P0498A12.6"  
/note="hypothetical protein  
similar to Oryza sativa chromosome 1, P0410E03.9"  
/codon\_start=1  
/protein\_id="BAB39978.1"  
/db\_xref="GI:13486744"  
/translation="MRPIFFPPEPRPRAAAYARDNAEGTKELIAPELPHFEBA  
AAATVYDREESDEEPTVIGTGLMSDFQYIGELGGGLKNYYTAKRLRPRPGVYAL  
KTPYVDIGREDEDAVAALVRVGLSHVYRCAMPRNRESLRAVAEHNNGSLDRA  
LSRGGRGSLPEPALAEVAACTALRGLHSRGVHLHNPFHIIADVDGNIKICDDPD  
AKIIPGPGSTVYSGIGFSPSEYMSPERCAMPAMGAAVAPAAADWAGLITVLAHOG  
YCPARLPDEKRTLERLSYEISGEPSPVEDPRASIEIRGFVACQKQCVCRATVA  
ELCNHPFAVRDVAASRVLKEVIVATMDQV"  
join(29480..30000,30031..30083,30109..30266)  
/gene="P0498A12.7"  
join(29480..30000,30031..30083,30109..30266)  
/gene="P0498A12.7"  
/note="hypothetical protein"  
/codon\_start=1  
/protein\_id="BAB39979.1"  
/db\_xref="GI:13486745"  
/translation="WSTFLHFFTLPLPSPSEFSLSLPSLPLVTPDQRRNRRCGGGG  
GVAATACGRGEMLRRAAGNRKACQPFQPLSPSELISLVASRRSNLITLGS  
TDEEQDEADGADPALVDPVDDSEKSVLLAAAEAPRNRMVQMVAATGSPVAVG  
FGESFTQVEARRKASLIARRLMSPGRSHSTLYKSPNNGRIRYHSSAAEELPADEL  
VATGPSAADAVRRRRESAAGALOK"  
complement(join(31539..32068,32196..32316))  
/gene="P0498A12.8"  
complement(join(31539..32068,32196..32316))  
CDS

/gene="P0498A12.8"  
/note="contains ESTs AU101991(S4037), AU070167(R0031)"  
/codon\_start=1  
/product="probable germin protein 4"  
/protein\_id="BAB39980.1"  
/db\_xref="GI:13486746"  
/translation="MARVQLWVAACAVLALAPSLADPDMLODYCVADLAPKLN  
NGPCKANTADDEPFAKNGNNPAGSNNTANQSPGVNLSVMAIDYAP  
GGQNPHTHRPARETFFVLEFVETITANTLTKTKYTAGEPVPRDLVHQQNR  
CHGPAAVIAFNSQGIQAIATLFAAPVPSDVLAKEKRVDPYDAIKAKER"  
gene  
misc\_feature  
38583..40973  
/gene="P0498A12.9"  
38583..40973  
/note="P0498A12.9"  
/gene="probably inactive due to stop codon(s) in CDS  
pseudogene, RIKO protein"  
gene  
join(43590..43652,44276..44830,44901..45063,45238..45768,  
49084..49165,49254..49427,49521..49631,49733..49805,  
49901..49999,50099..50278)  
/gene="P0498A12.10"  
join(43590..43652,44276..44830,44901..45063,45238..45768,  
CDS  
Query Match 2.98; Score 177.8; DB 8; Length 159749;  
Best Local Similarity 83.3%; Pred. No. 8.7e-27;  
Matches 214; Conservative 0; Mismatches 42; Indels 1; Gaps 1;  
ORIGIN  
1967 GGGCGATCCAGAACAAAGTTGGGGGATCATCATCCAGTACATGCAATTAAC 2026  
|||||  
Db 107392 GGGCGATCCAGAACAAAGTTGGGGGATCATCATCCAGTACATGCAATTAAC 107451  
|||||  
QY 2027 ACATCATCTCAGTATTAAGTATGCTAAATGCTATTAAGAGACCTTTAGCACTCTAT 2086  
|||||  
Db 107452 AATATATCTCAATATTTAGTACTATGCTAAATGCTATTAAGAGACCTTTAGCACTCTAT 107511  
|||||  
QY 2087 CTATCAACCATGAGGAAATTAAGAGGAGGAGCTCAGGAGGCTTCCATGGGCGAT 2146  
|||||  
Db 107512 CTATCAACCATGAGGAAATTAAGAGGAGGAGCTTCCATGGGCGAT 107571  
|||||  
QY 2147 GGGTGCAGGGGGGAGC-TGAGTCCGCCCTGCACCCAGCTTGAATCCGCCCTGGCATGGCTA 2205  
|||||  
Db 107572 GGGTGCAGGGGGGAGC-TGAGTCCGCCCTGCACCCAGCTTGAATCCGCCCTGGCATGGCTA 107631  
|||||  
QY 2206 TAAGCTGTCAAGCCAT 2222  
|||||  
Db 107632 TGACCTGTCAAGCTTCAAT 107648  
|||||  
RESULT 14  
AC134928 152243 bp DNA linear HNG 03-OCT-2002  
LOCUS Oryza sativa (japonica cultivar-group) chromosome 5 clone B1164G01,  
DEFINITION \*\*\* SEQUENCING IN PROGRESS \*\*\*; 3 ordered pieces.  
ACCESSION AC134928  
VERSION AC134928.1 GI:23477777  
KEYWORDS HNG; HTGS\_PHASE2.  
SOURCE Oryza sativa (japonica cultivar-group)  
ORGANISM Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Eriharoidae; Oryzaceae; Oryza.  
1 (bases 1 to 152243)  
REFERENCE  
AUTHORS Chow,T.-Y., Hsing,Y.-I.C., Chen,C.-S., Chen,H.-H., Liu,S.-M.,  
Chao,Y.-T., Chang,S.-J., Chen,H.-C., Chen,S.-Y., Hsiao,S.-H.,  
Chen,Y.-L., Cheng,C.-H., Chung,C.-I., Han,S.-Y., Hsiao,S.-H.,  
Hsiung,J.-N., Hsu,C.-H., Huang,J.-J., Kau,P.-I., Lee,M.-C.,  
Luo,H.-L., Li,Y.-F., Lin,S.-J., Lin,T.-C., Wu,S.-W., Yu,C.-Y.,  
Yu,S.-W., Wu,H.-P., and Shaw,J.-F.  
Oryza sativa BAC B1164G01 genomic sequence  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 152243)  
AUTHORS Chow,T.-Y. and Hsing,Y.-I.C.  
TITLE Direct Submission

JOURNAL Submitted (03-OCT-2002) Institute of Botany, Academia Sinica, 128,  
Section 2, Academia Road, Nankang, Taipei 11529, Taiwan  
COMMENT  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 3 contigs. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have  
\* provided by the submitter.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
\* 1  
\* 729: contig of 729 bp in length  
\* 829: gap of unknown length  
\* 830 146207: contig of 145378 bp in length  
\* 146208 146307: gap of unknown length  
\* 146308 152243: contig of 5936 bp in length.  
FEATURES  
source  
1..152243  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="genomic DNA"  
/cultiar="Nipponbare"  
/db\_xref="taxon:39947"  
/chromosome="5"  
/clone="B1164G01"  
/note="(japonica cultivar-group)"  
BASE COUNT 45337 a 31135 c 30898 g 44673 t 200 others  
ORIGIN  
Query Match 2.98; Score 173.4; DB 2; Length 152243;  
Best Local Similarity 78.1%; Pred. No. 7.5e-26;  
Matches 221; Conservative 0; Mismatches 61; Indels 1; Gaps 1;  
BASE COUNT 45337 a 31135 c 30898 g 44673 t 200 others  
ORIGIN  
1957 GATTGATTTGGGGGATCCAGAACAAAGTTGGGGGATCATCATCCAGTACATGCAATTAAC 2016  
|||||  
Db 23014 GTGCTTCAGAGACCGATCCAGAACAAAGTTGGGGGATCATCATCCAGTACATGCAATTAAC 23073  
|||||  
QY 2017 TGGGTAACACATCATCTCAGTATTAAGTATGCTAAATGCTATTAAGAGACCTTTAG 2076  
|||||  
Db 23074 CGATTAACACATCTCAGTATTAAGTATGCTAAATGCTATTAAGAGACCTTTAG 23133  
|||||  
QY 2077 CACCTCTTATCTTATCAACCATGCTGAATAAATTAAGAGGAGGAGCTCAGGAGGCTATCA 2136  
|||||  
Db 23134 CACCTCTTATCTTATCAACCATGCTGAATAAATTAAGAGGAGGAGCTCAGGAGGCTATCA 23193  
|||||  
QY 2137 TGGGTCGATGGGTGAGGGGGGAGC-TGAGTCCGCCCTGCACCCAGCTTGAATCCGCCCT 2195  
|||||  
Db 23194 TGGGTCGATGGGTGAGGGGGGAGC-TGAGTCCGCCCTGCACCCAGCTTGAATCCGCCCT 23253  
|||||  
QY 2196 GGCATGCGTATTAAGCTGTCAACGCAATTTAGAGTCTTGTGC 2238  
|||||  
Db 22354 TGCGGTCTTATTAATTAAGAACACTTAATAAAGTAGTGTGC 23296  
|||||  
RESULT 15  
OSJUN0064 121799 bp DNA linear PLN 12-JUL-2002  
LOCUS Oryza sativa genomic DNA, chromosome 4, BAC clone: OSJUN0085K24,  
DEFINITION complete sequence.  
ACCESSION AL606609.2 GI:21740604  
VERSION AL606609.2 GI:21740604  
KEYWORDS HNG.  
SOURCE Oryza sativa (japonica cultivar-group)  
ORGANISM Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Eriharoidae; Oryzaceae; Oryza.  
1  
Fu,G., Wang,S.Y., Ren,S.X., Lv,G., Lin,W., Gu,M.Q., Zhu,G.F.,  
Tu,Y.F., Jia,J., Yin,H.F., Zhang,Y., Cai,Z., Chen,J., Kang,H.,  
Chen,X.Y., Shao,C.Y., Sun,Y., Hu,Q.P., Zhang,X.L., Zhang,W.,  
Wang,L.J., Ding,C.W., Sheng,H.H., Gu,D.L., Chen,S.T., Ni,L.,  
Zhu,F.H., Han,B., Feng,Q., Huang,Y.C., Li,Y., Zhu,J.J., Zhao,Q.,  
Hu,X., Liu,Y.L., Mu,J., Yu,Z., Chen,L., Fan,D.L., Weng,Q.J.,



TITLE  
 JOURNAL  
 Submitted (27-JUN-2002) Han Bin, National Center for Gene Research,  
 Chinese Academy of sciences, 500# Cao Bao Road, Shanghai 200233,  
 CHINA. E-mail enquiries: bhan@ncgr.ac.cn. Clone requests:  
 bhan@ncgr.ac.cn  
 REMARK  
 Oryza sativa japonica (nippondare) genomic DNA, chromosome 4, BAC  
 clone: OSJNB0089K24.  
 On Jul 12, 2002 this sequence version replaced gi.15594068.  
 COMMENT  
 Web site: <http://www.ncgr.ac.cn/>  
 ----- Summary Statistics -----  
 Assembly program: phrap  
 -----  
 NOTE: This is a complete sequence.  
 Genes were identified by a combination of several methods: Gene  
 prediction programs including Egenesh (<http://www.softberry.com/>),  
 genscan (<http://CCR-081.mit.edu/GENSCAN.html>), GenemarkEM  
 (<http://genemark.biology.gatech.edu/genemark/>), tRNAscan-SE (Sean  
 Rddy, <http://genome.wustl.edu/eddy/tRNAscan-SE/>), searches of the  
 complete sequence against NCBI none redundant protein database (nr)  
 (<ftp://ncbi.nlm.nih.gov/blast/db>) and the EST database at NCGR.  
 Location/Qualifiers  
 1. 121799  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /mol\_type="genomic DNA"  
 /variety="Nippondare"  
 /sub\_species="japonica"  
 /db\_xref="taxon:39947"  
 /chromosome="4"  
 /clone="OSJNB0089K24"  
 /clone\_1id="CG1-OSJNB"  
 complement(3496..6534)  
 /gene="OSJNB0089K24.1"  
 complement(join(3496..3796,3876..4046,4129..4961,  
 5043..5504,5583..6316,6405..6534))  
 /gene="OSJNB0089K24.1"  
 /codon\_start=1  
 /protein\_id="CAD39392.1"  
 /db\_xref="GI:21740605"  
 /translation="MSTPGANPSPPTPARIOEANPATEAVAGHPATTAVOQEIHE  
 LPDOHSMGSGTSASRSSRNRSNIMPTVOYIRQVADSGRPAPRTVYGRSMSC  
 GIARERNGLIKDQIKYVEAEKBRAPAEKPTTPAPAKRKRKRLKQMKAKN  
 WSKLTFEYVNPGRHTPPDEIPQITRAVEEPCSKITTOERESSEBHRULQKNEH  
 PRLGTAGYIGKEALVAQDAALAAANPAPFSDIPEQARWARAKGVNPDGVTF  
 ENKSDAVYQELGLVAQASQSEVESPAPKRREDIITKALGTKEHPRTGIGSDVP  
 WKHGLPOYSQYRRKRYKEERDAIKAEKVEYIOELASMANAVERKAVLADN  
 IPRYTPAVOQPTRIQHDASPEOHSSCASAEVAPAGLPIDPLAVADIEGAQVTL  
 ARVPTAPAVAGMAKRPSTVDYKAGADILAGAVSIDTYKDWGSPILVPSPDE  
 ITTIGDAKPTFIOWPREKEDIVYMPRPSPRPSPKSKTSLTEARSPALSVPSGQ  
 ADMDLADIAOSLAPITTTTKADSSPPLYKQKREKRGVGLAEPRKRAATSNPY  
 SKSGVVRAPAEFELGPDVLEDVNAVGLVWGLRCHEKLYMELSNKAKRRSSIVGH  
 DHQPLSSPAYITIGPDLFDLFRILKIDGLKCYSLCLWIESRHNQVGLDPSM  
 VNEVLAQSSTEVDYVNRCLMAHQDEKTYICAHQNERHMLLVIVPMSVYTLN  
 KSKYDSEITKALNMAWGPYVEKGGKRRKEGKNELVHDKFACAOQIGDQGFVCHN  
 MSTLREVKDPEEVANGE"  
 complement(9689..12976)  
 /gene="OSJNB0089K24.2"  
 complement(9689..12976)  
 /gene="OSJNB0089K24.2"  
 /codon\_start=1  
 /protein\_id="CAD39393.1"  
 /db\_xref="GI:21740606"  
 /translation="MDRMMATYAHRSSTYRBEVTEVTPPANDRKSRSKSHMLPCR  
 DCKNEQMIENKDEYVAHLIMNGFNKRYCWTKEHQEQPADVAEEVLDQVETAAAR  
 EGMFVPSLGGETIDIDQCLSTMLDHIIDAENDRDYKFSKLVEDQMLDQSGS  
 KHSKISCVLEMLKASNGSDKSFTELETLKDLPEGNLIPQITYAKVYLCPLI,  
 EVYRIHACPDQCLITYYKVAADVCPICGAEKRSKSGSGSKSGGAPKVVYTP  
 IAEKRIKMFANKDQAKLVMAHEERKVDYTMRLRPAASVOMRTIDRIYGFENDRNMR  
 PAKCTGGINPGDLSRHSSTPULLVYNLPPMLCPKRYIMLALITGPPQPDND  
 VFLEPIIDDEFRLMNGEITMDIAQETFNLEAPLCTINDYFALGNLSGQTVAGKMA

CSCSCEBETRSKVLHSHKTYVMGRHEFLPRHPTPRNNKRNPNHEDTAPPELTGTE  
 VHLVWGTINTEFGKRRVVGKRRKSTYSKEIEHEKQKREKSMKKSITWRLPW  
 KDLVEVRIIDIMAEKNCESLMLGDLNPNRKGALNARDIDMKNVETLPTTES  
 GRVYLPACAYTLSEKEKIDLTLCISIGIKRSGSISRLVYSIDQKLVGMKSHCHV  
 LITOLLPAVINLPPRYKRTIORLCSFHAIAGOKIIDPDELDELQETLCTHEM  
 YPEPPPIPMHLPVNHYROTCKCGAEPQWQYQCEYRITLGRYRNSHPESSITL  
 STTEBAPLBFQVNDSETSSTGLPRSHHBRGLDYGIVGKRTIRLRYKTYKAAHTYL  
 QHTEVVPYVDEHLAVIRQNPGRSPSRNKNSSSEFMKLNRIARLQNLSETLQW  
 LSGQEWASATWQGDINGYTFPTVKQDSKQVNSGRIEASDGRDRDYGVQEO  
 ILDELITAEKYDLFCRMWDLIRNWKYDNAPFTYVLANAKDEPYLAKOVQVLYI  
 VDCNKLHVVREGRRRTVGLNDLADDDYQNHVAGIQGELPLEEEDDEYQVARNV  
 HEEGTEFL"  
 15539..27175  
 /gene="OSJNB0089K24.3"  
 join(15539..15646,16171..16262,17229..17326,17436..17591,  
 22965..23578,23692..23762,24199..24444,25233..25359,  
 25959..26151,26622..27175)  
 /gene="OSJNB0089K24.3"  
 /codon\_start=1  
 /protein\_id="CAD39394.1"  
 /db\_xref="GI:21740607"  
 /translation="MSKLSRQSPFLFLLQFLDLPWSSNALXGEPPTKLYTYLGE  
 RRHDDADLVTDSDHMDLASYLSGKEALDSIYSYRFSQFARLTKAQAASIRLPL  
 DYVSVRENIHQLERSKEMPELMDYRQNGLLAANYGEDIIGVLDGTIPESHSF  
 ADGVPPEPSKKKGIQVQPSPEAKGONRKLIGAWYIDDDTSLMSKNEILSPDVE  
 GCHGTASAGNITVHNASTILGLATGVAGCAPRARVAMYITCMSSGCGSAAYCAL  
 DDAVYDGVDSLSLGSPELDLITLHVAKGIPIYYSKNGDGPVQIYEMSSKRLTY  
 AAATMKRSPFVYITLGDNHFVAAQSFVLSRQTSFSETIOFEHDDCANLINSYVG  
 KIVFCPTGLDPEPDINSIITKVGEGGTVGIMPRTNDITLDQGLTLPPIPVVDY  
 ELAVRYQYTNENDGAKVKSISITQITIGKTAKVAAFSRSPSSIVGKEDIA  
 AVQVTLAAPKNVNDIGLIPYHPSGTSMACPVSGVIAILKSHLPMSPAAKLSIM  
 TTAITYDNDGQIOANGVOKIADPDYAGFRLADPDYADGLIYDISADYKLPKNC  
 MGLSGDCKCTTYKGSADLNLPSISIPMLKTIQATYATYVNGQANAYAKALDPPV  
 GIMAAVEPMLVFSKDRKVOSEFVYTRVIRPFIQSDYFSGSLAHWDGHWVRIPLAV  
 RIVIEIYSKIS"  
 33167..36845  
 /gene="OSJNB0089K24.4"  
 join(33167..33498,35578..35645,36790..36845)  
 /gene="OSJNB0089K24.4"  
 /codon\_start=1  
 /protein\_id="CAD39395.1"  
 /db\_xref="GI:21740608"  
 /translation="MAADPGLIYIEPSDYFFNOMGSLGRDNTYKESYADNLT  
 PSVAIKRLKTSNKRVSYTNVDYQANAYAKALDPPYGVMAVDVSIVFSEEEKVLSF  
 KVAIKMRKKSINKMDKVIETRIEIKLTKSSKNGPQNNLAAVLTATKVA"  
 47014..51905  
 /gene="OSJNB0089K24.5"  
 join(47014..47115,47730..47821,48103..48200,48299..48454,  
 48702..49321,49446..49513,49814..50065,50356..50491,  
 51020..51212,51352..51905)  
 /gene="OSJNB0089K24.5"  
 /codon\_start=1  
 /protein\_id="CAD39396.1"  
 /db\_xref="GI:21740609"  
 /translation="VAPRSKIALFLVLLPQNLQDLSYGLQDQSRILYTYLGRQ  
 HEADADIVSHDMILTSILSGKEETLASIYSTRGFGFSAMLTQSQARKILGELV  
 LSVYEMQITKHTSRMDPLGIDKTPNGLAKARGDGLIIGVDDTIPESPSFD  
 AGCGPPSKWKGIQVQPSFGTNCNKRITGARWAVYNDGTIDTDELPDYHSHG  
 TRASRAGNIVAHNSRLGIAGTAGGAPRALAYVACATATDGTGCGGAGLITAM  
 DDRIHGVDTLSISIGSPEEMGTIHYANGILVYISAGNCPINQYBMSRWLLTY  
 AAAMKSPFVYITLGNNEKFAVQSFVYSSASQSESTMDNDMCADNDITVAGM  
 IVFETIKEDMENYDRILINTVASYASVARGSGVIFPKSTQDILFREDLITFIPVLY  
 DYEISYIRQYLIINNENGINIPKASISLTKTMVGSNSAPKILAAKSLHPEMSPALK  
 PDLAAGVAILIASSNPTEPFKGVYRFDSGTSACPHVSGIILAYLKLHPEMSPALK  
 SAITMTANTFQNNKGMQANGRPVKIADPDYAGVYNDMAADPGLIYIINPDLTK  
 FRCMGSLSGQNCCTTYGSIYDNLPSLIPMLKSEIYATYVNTVNGQOEVYVYAF  
 LDPPAGLNAVSESELVSKDKQDSFVTFKATRVQDGTIPGSLAHMGSGSHWRI  
 PIAVHYIEIYSNIS"  
 57504..61838  
 /gene="OSJNB0089K24.6"  
 join(57504..57555,57594..58598,58647..59953,59992..60966,  
 61080..61838)  
 /gene="OSJNB0089K24.6"

gene

gene

CDS

gene

CDS

gene

CDS

gene

CDS

```
/codon_start=1
/protein_id="CAD39397.1"
/db_xref="GI:21740610"
/translation="MSFVISTPLINNNVVIDLLHOMQNPQAPVQAPIHAPQSKLS
EELVRRPPTFSSTNPVEANDMLHAVNRKIDLIQCTDEKVAFAHQLHGPASEMDH
FOATQAGOVITWACESTAFRRRAHPAGVVALKKEFEELKQDGRVTEFLAEFRFA
RYADEVPTDEEKOEFLEGNDELAVOLISGDYEPQKLVDKALILEDKRRMESEK
HKMANAMFQPSOKPRVPEPPVOGSSSTPAORAPRQNNKLQNSGNNQORSTPAOG
GSTPKDSCKIIVCYNCEPGHGFADKCPKPRQAPAPAHNANOSNNNNNNNLYRG
RVNHVSEBAQAAPVYITIGATHSTSDPFAKHQIRNEDLSKPMITIPSGQLPAKL
SPSVTLEIGVTFELANLILISSKNLDVILGMNLTNHSGLIDCASRISLIGTGQRV
IFHAKPALPKQVLAQVMSMEIPIVKEIPDVPEELPGMPKRDIEFRIDIVPGTAPI
SKRYTMAANELAIEVKOLEKIKGKHIRPSTSPMGAPVIFVEKDKTRKMCYDVAL
NEVTKKKYPLPRIDLPNOLKATFVSKIDLSGYHOLRIEODIPKTAFTTGYLY
ECYMSRGLTNAPAFPMNLANKVFMEYVIFIDOLITYSKTAPEHROHCLALE
KLKXHQLIYAKESCEPWTSEVKTIQHTISQAGANDPVVESVTNWDOSKSYTVRSF
LGLAGYTRRPIENFSRIARPMTEILKEKEFWTQCEAPQLKRLDFQVYTDASK
QGLGCVLMQGOVVAASRQLRPHESNYPHDELAAVVAHAKIMQYLIGNHCELYT
DHKSLKTIETQDNLIRQRMLEF IKDYDGLIHYHIGKANVADLSRKSVCNAIIE
OCPPGLCRELDALNGLIYSYGVATLEAPPLIDHIREAKGSDSDIOEIKNNRGRKA
IGFSEDEHGTIMYGRACVPNDPEIKRLILOEASHSOYSIHPSGTMVODLKKMYMW
SMRETIPEPALODVCORVKAHORPAGILOPKIPEKMEIEMDPIITGLPRTNDM
TTEGLISTAKPKLHISOSKPLIPAEKTHLNFSTAHHPQDQOTERNQIILEDLR
ACAMDRKSWDSLPAEFSYNSYSQASIQWAPFEALIGKRCRPLMDQTERQVFG
TDIIEAEKVKVIOANLKAOSRQSKYATRRRELAEGDFYLYKTVPLRGVRPFT
KGMLEPFVQPYVIGRGEVYOLEPASTIGVYDFHVSQKKYLRVPTQADPEQ
IKIQEDLYVEKPIRIETSERHTRSVIRFCVQWSNHLREAF"
complement(62678..67575)
/gene="OSUNB0089K24.7"
/complement(join(62678..64405,64756..65143,65489..67575))
/cds
/gene="OSUNB0089K24.7"
/protein_id="CAD39398.1"
/db_xref="GI:21740611"
/codon_start=1
/protein_id="CAD39398.1"
/translato="MTGNNBAVATNOCGRPKSNPAETNGASTSELEKSDAARPC
PFDNHESTMTSEATRSYISIKTRKHTILOACWFLNRAIRACKERGIQRIPTPR
```

```
Query Match      2.88; Score 172.2; DB 8; Length 121799;
Best Local Similarity 87.6%; Pred. No. 1.3e-25;
Matches 211; Conservative 0; Mismatches 28; Indels 2; Gaps 2;

QY 1967 GGGGGATCCAGAGCAAAAGTTGGGGGATTCACATACGAGTACACGATGCAATAAC 2026
DB 77946 GGGGGATCCAGAGCAAAAGTTGGGGGACTCACATACGAGTACACGATTAAT 78005

QY 2027 ACATCATCTCAGATTAACTATGCTAAAGCTATTAGAGACCTTTAGACCTTTAT 2086
DB 78006 ACATATCTCAGATTAACTATGCTAAAGCTATTAGAGACCTTTAGACCTCTAC 78065

QY 2087 CTTATCAACCATGCTGTAAGGAGGAGCTCAGGGGGGTATCCATGGGTCCGAT 2146
DB 78066 CTTATCAACCTATGATGAAAATGTAAGGGGGGCTTA-GGGGGGTATCCATGGGTCTTGT 78124

QY 2147 GGGTGCAGGGGGGAC-TGAGTCCCCCTGCACCCACGTTGATCCGCTGCGCATGCGTA 2205
DB 78125 GGGTGCAGGGGGGACTGAGTCCCCCTGCACCCACGTTGATCCGCTGCGATGTAATC 78184

QY 2206 T 2206
DB 78185 T 78185
```

Search completed: September 13, 2003, 23:42:52  
Job time : 14947.6 secs



XX Producing heterologous polypeptide in plant grain, by culturing  
PT transformed plant to form a grain-producing transforming plant, and  
PT recovering transgenic grains containing polypeptide from transformed  
PT plant  
PS Disclousure; Fig 42A-I: 230pp; English.

CC The invention describes a method of producing a heterologous polypeptide  
CC (I) in a grain of a plant, comprising culturing a transformed plant (P1)  
CC comprising a first chimeric gene, and optionally, at least one  
CC heterologous transcription factor that is capable of enhancing the  
CC expression of the chimeric gene, to form a grain producing transforming  
CC plant (P2), and recovering transgenic grains containing (I) from P2.  
CC The method is useful for producing heterologous polypeptide in a grain  
CC of a plant. (I) is a non-plant storage, human or non-human animal, milk  
CC or other than a milk polypeptide, antibodies, cytokines, lymphokines,  
CC chemokines, hormones, growth factors, coagulation factors,  
CC anti-infectives, or cytotoxins, or anti-inflammatory molecule or  
CC intestinal trefoil factor (ITF) or its active fragment. Preferably, (I)  
CC is lactoferrin, lysozyme, lactoferricin, ITF, epidermal growth factor  
CC (EGF), keratinocyte growth factor (KGF), insulin-like growth factor I  
CC (IGF-I), lactoperoxidase, kappa-casein, haptocorrin, lactoperoxidase,  
CC beta-1-antitrypsin, immunoglobulins, alpha-lactalbumin, fibrinogen or  
CC beta-lactoglobulin, alpha-casein, beta-casein, albumin, fibrinogen or  
CC protease inhibitor. This sequence represents a polynucleotide associated  
CC with method of producing a transgenic plant.

SQ Sequence 6227 BP; 1683 A; 1312 C; 1400 G; 1832 T; 0 other;

Query Match	100.0%	Score 6227	DB 24	Length 6227
Best Local Similarity	100.0%	Pred. No. 0		
Matches 6227	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY 1	GGTACCATCTAATACATTAATACAGAGAGAGATGATATGCAATTTATTTT	60		
DB 1	GGTACCATCTAATACATTAATACAGAGAGATGATATGCAATTTATTTT	60		
QY 61	ATGGAGGCTATTTTATGCGATTTTGTAAATACAGGCGCAATTCCTACTAGTA	120		
DB 61	ATGGAGGCTATTTTATGCGATTTTGTAAATACAGGCGCAATTCCTACTAGTA	120		
QY 121	AAGCTACGATGACTATGCTACCGCTACGAGTGTGATTTGGAATTCCTGATGATC	180		
DB 121	AAGCTACGATGACTATGCTACCGCTACGAGTGTGATTTGGAATTCCTGATGATC	180		
QY 181	TGTTGCTGTTGCGAGTTAGGTTAGTCTGATACCTCCAGCGGAAACAGAGGCTTTC	240		
DB 181	TGTTGCTGTTGCGAGTTAGGTTAGTCTGATACCTCCAGCGGAAACAGAGGCTTTC	240		
QY 241	GCAAGTTTATGAGTGTCCAAAGTTAGCTAGGCGCAATTCATGTTCAAGTATTTGAG	300		
DB 241	GCAAGTTTATGAGTGTCCAAAGTTAGCTAGGCGCAATTCATGTTCAAGTATTTGAG	300		
QY 301	TATATGAAAAAGAGATCTCCAAACATTTAATTTGTATTAAGGAGAAATCGAATC	360		
DB 301	TATATGAAAAAGAGATCTCCAAACATTTAATTTGTATTAAGGAGAAATCGAATC	360		
QY 361	TGAGGTCTTAATTCACACGAGCTACCTCCCTCGTTCAATTAATGATTAATACATA	420		
DB 361	TGAGGTCTTAATTCACACGAGCTACCTCCCTCGTTCAATTAATGATTAATACATA	420		
QY 421	TACGATATATGATATACACATATATATATATATATATATATATATATATATAT	480		
DB 421	TACGATATATGATATATACACATATATATATATATATATATATATATATATAT	480		
QY 481	AT	540		
DB 481	AT	540		
QY 541	TTAT	600		
DB 541	TTAT	600		

QY 601	AGTATGAGATGCAAT	660		
DB 601	AGTATGAGATGCAAT	660		
QY 661	TAAAGCAT	720		
DB 661	TAAAGCAT	720		
QY 721	CTAATTCGCTGAT	780		
DB 721	CTAATTCGCTGAT	780		
QY 781	TATCTCTATGAGGCTATGAGGCTATGAGGCTATGAGGCTATGAGGCTATGAGG	840		
DB 781	TATCTCTATGAGGCTATGAGGCTATGAGGCTATGAGGCTATGAGGCTATGAGG	840		
QY 841	TTTTGCGGTATGAT	900		
DB 841	TTTTGCGGTATGAT	900		
QY 901	AATTTCCGATGAGGCTATGAGGCTATGAGGCTATGAGGCTATGAGGCTATGAGG	960		
DB 901	AATTTCCGATGAGGCTATGAGGCTATGAGGCTATGAGGCTATGAGGCTATGAGG	960		
QY 961	TCCGATGAT	1020		
DB 961	TCCGATGAT	1020		
QY 1021	TCCGATGAT	1080		
DB 1021	TCCGATGAT	1080		
QY 1081	AATCTGAT	1140		
DB 1081	AATCTGAT	1140		
QY 1141	GAGAGGATGAT	1200		
DB 1141	GAGAGGATGAT	1200		
QY 1201	ATTCGATGAT	1260		
DB 1201	ATTCGATGAT	1260		
QY 1261	CGGATGAT	1320		
DB 1261	CGGATGAT	1320		
QY 1321	GAGAGGATGAT	1380		
DB 1321	GAGAGGATGAT	1380		
QY 1381	CCGATGAT	1440		
DB 1381	CCGATGAT	1440		
QY 1441	GAGAGGATGAT	1500		
DB 1441	GAGAGGATGAT	1500		
QY 1501	GAGAGGATGAT	1560		
DB 1501	GAGAGGATGAT	1560		
QY 1561	GAGAGGATGAT	1620		
DB 1561	GAGAGGATGAT	1620		
QY 1621	GAGAGGATGAT	1680		
DB 1621	GAGAGGATGAT	1680		

QY	1681	AGACCAATGGCGTCACATGGGTCTTGCAATGGCGTCGTGTGATGTGCGTGGTGGTGGTGCAC	1740
Dp	1681	AGACCAATGGCGTCACATGGGTCTTGCAATGGCGTCGTGTGATGTGCGTGGTGGTGGTGCAC	1740
QY	1741	TCATAGATGTAGAGGTGGCGTATGACAAAGGAGAGCTTCATAGACCGCGATATAG	1800
Dp	1741	TCATAGATGTAGAGGTGGCGTATGACAAAGGAGAGCTTCATAGACCGCGATATAG	1800
QY	1801	ATTGGCATTCGATGTCACATCTGCGAATTTGGCGATGTAACTGATCAATTCATCTAAACT	1860
Dp	1801	ATTGGCATTCGATGTCACATCTGCGAATTTGGCGATGTAACTGATCAATTCATCTAAACT	1860
QY	1861	GCGAGACAGTGGTCTTTGCAATTTGCAATTTTGTATGGGCGCTTAAACGTGAT	1920
Dp	1861	GCGAGACAGTGGTCTTTGCAATTTGCAATTTTGTATGGGCGCTTAAACGTGAT	1920
QY	1921	TGCGTTTATAGATTGGCAATATGTGACTTTATGCAAGTATTTGATTTGGCGGATCCAGGA	1980
Dp	1921	TGCGTTTATAGATTGGCAATATGTGACTTTATGCAAGTATTTGATTTGGCGGATCCAGGA	1980
QY	1981	NCAAAAAGTTGGGGGGGATTCACACTACCGAGTTCACGTGGCATTAACACATCATCTCAGTA	2040
Dp	1981	NCAAAAAGTTGGGGGGGATTCACACTACCGAGTTCACGTGGCATTAACACATCATCTCAGTA	2040
QY	2041	TTAAACTATCTCTAAATGCTATTAAGAGACCTTTAGACACCTTATCTTATTCACACATGG	2100
Dp	2041	TTAAACTATCTCTAAATGCTATTAAGAGACCTTTAGACACCTTATCTTATTCACACATGG	2100
QY	2101	TGAAAAAATTTGAAAGGGGGGACTCAGGGGGGTATCCATGSGGTCCAGTGGGTGTCAGGGGGA	2160
Dp	2101	TGAAAAAATTTGAAAGGGGGGACTCAGGGGGGTATCCATGSGGTCCAGTGGGTGTCAGGGGGA	2160
QY	2161	CGAGTCCCCCGACACCCAGGTGGAATCCCGCCCGGACGACGATAGCTAGCTGTCACAGCC	2220
Dp	2161	CGAGTCCCCCGACACCCAGGTGGAATCCCGCCCGGACGACGATAGCTAGCTGTCACAGCC	2220
QY	2221	ATTTCAGTGTCTGTGCTAGTAAATGGTGGTATGTCACTTAAATTTGCTTTTCTATGTGT	2280
Dp	2221	ATTTCAGTGTCTGTGCTAGTAAATGGTGGTATGTCACTTAAATTTGCTTTTCTATGTGT	2280
QY	2281	CATGATTTTCTAAGAAAGSAAAAATAGCCATTTATTTAGTTCAGAAATTTGATATGCC	2340
Dp	2281	CATGATTTTCTAAGAAAGSAAAAATAGCCATTTATTTAGTTCAGAAATTTGATATGCC	2340
QY	2341	TGGCCCTTCATTTTGGCAAAATTAGCCATTTGTGGTGGCCCTTCAGATTTTCCAGGT	2400
Dp	2341	TGGCCCTTCATTTTGGCAAAATTAGCCATTTGTGGTGGCCCTTCAGATTTTCCAGGT	2400
QY	2401	TATGTGTGTGCAATGTGTGGCTATGGCTCGTTTGTGGCCATATATTTATTTATTTGCAAT	2460
Dp	2401	TATGTGTGTGCAATGTGTGGCTATGGCTCGTTTGTGGCCATATATTTATTTATTTGCAAT	2460
QY	2461	TCATTTTGTGACATGACTTAAATAGACATAGAGCAACATGCACTGATGTGGTATTCAT	2520
Dp	2461	TCATTTTGTGACATGACTTAAATAGACATAGAGCAACATGCACTGATGTGGTATTCAT	2520
QY	2521	TAATCATTTATGTCTGTCTGTTCATTTATATCTAGTACTGATGTCATTTTCATCTTCAGG	2580
Dp	2521	TAATCATTTATGTCTGTCTGTTCATTTATATCTAGTACTGATGTCATTTTCATCTTCAGG	2580
QY	2581	CGTCGGACAGTTCACACCTGAGCGTCTGGACCTGTTCATCCTTCTGTATAGCAGATG	2640
Dp	2581	CGTCGGACAGTTCACACCTGAGCGTCTGGACCTGTTCATCCTTCTGTATAGCAGATG	2640
QY	2641	TTTCCACATAGCGCGCTCTAATTCATGCGAGGATCTATCTTATCTGCTTTCATATTTTC	2700
Dp	2641	TTTCCACATAGCGCGCTCTAATTCATGCGAGGATCTATCTTATCTGCTTTCATATTTTC	2700
QY	2701	AGATTTTATGAAACACACCAAAATATCTGCAAAATGTAGAGGATTAATTTTATAGTTC	2760
Dp	2701	AGATTTTATGAAACACACCAAAATATCTGCAAAATGTAGAGGATTAATTTTATAGTTC	2760
QY	2761	TTTGAATAAGATAGAGTTGATTTGTGTGTCACGTATCATCAATATAGTTCTTACAGCTATGA	2820

Db	2761	TTTAAATTAAGATAGAGTTGATTTGCTGTCAAGCATCAATATAGTTCTGAGAGCTATGAA	2820
OY	2821	TAAATAGTCCGCAATTTGATAGATCTTTGACATTAATATGTTATGCTTAAGT	2880
Db	2821	TAAATAGTCCGCAATTTGATAGATCTTTGACATTAATATGTTATGCTTAAGT	2880
OY	2881	ATAGAGTTATGTTGTTGGAGTTCCTTAAATCAATTCATGCTGACTCCAGCTGGCA	2940
OY	2941	GGAGATTTGTGTGCTTGACATTAAGAGAGACCTTCCTGTTCTGAATGCTCAAG	3000
Db	2941	GGAGATTTGTGTGCTTGACATTAAGAGAGACCTTCCTGTTCTGAATGCTCAAG	3000
OY	3001	AAAACATTTTATTAATGACACCTTAATCCTTAGAGCTTGGCAAGAGAGGCACTT	3060
Db	3001	AAAACATTTTATTAATGACACCTTAATCCTTAGAGCTTGGCAAGAGAGGCACTT	3060
OY	3061	CTTTAGATTCGATAGTACTTAATACCTGTGTATCATGATTTGCTGTGTAGT	3120
Db	3061	CTTTAGATTCGATAGTACTTAATACCTGTGTATCATGATTTGCTGTGTAGT	3120
OY	3121	GTGCACAAACAGGATGCTACTTTTACCTTGATATTAATACCTGCTTACTAT	3180
Db	3121	GTGCACAAACAGGATGCTACTTTTACCTTGATATTAATACCTGCTTACTAT	3180
OY	3181	CTCTGTTTGGACATGGGACATATCAGGGCAATGCTACTCAGTTCAAAACATGCTA	3240
Db	3181	CTCTGTTTGGACATGGGACATATCAGGGCAATGCTACTCAGTTCAAAACATGCTA	3240
OY	3241	GGCCCAAGTGGGGAGTCGGGCTCAAGTGGTCAAGTGGATGATGATGATGATGATG	3300
Db	3241	GGCCCAAGTGGGGAGTCGGGCTCAAGTGGTCAAGTGGATGATGATGATGATGATG	3300
OY	3301	CCCAACAGCTCTTCATCAAGGAGAGAGTCAAGTATGATGATGATGATGATGATG	3360
Db	3301	CCCAACAGCTCTTCATCAAGGAGAGAGTCAAGTATGATGATGATGATGATGATG	3360
OY	3361	ACCACTGGACTCAAGACCTGTGTATCAAGATTTCAACAGGTGATATTCATTTGT	3420
Db	3361	ACCACTGGACTCAAGACCTGTGTATCAAGATTTCAACAGGTGATATTCATTTGT	3420
OY	3421	TCCTGTTAAATAGATTTGTACATTAATTAACCTGATGATGATGATGATGATG	3480
Db	3421	TCCTGTTAAATAGATTTGTACATTAATTAACCTGATGATGATGATGATGATG	3480
OY	3481	TTTTTACCTTAGACCCCTTAGTTGAGTTTCCACTTTCACCTGATGACCTTACCTT	3540
Db	3481	TTTTTACCTTAGACCCCTTAGTTGAGTTTCCACTTTCACCTGATGACCTTACCTT	3540
OY	3541	ATTGACAGCATATATGATCTTTACCTACCATATATAGTCTGTGGCCATACACAT	3600
Db	3541	ATTGACAGCATATATGATCTTTACCTACCATATATAGTCTGTGGCCATACACAT	3600
OY	3601	GCTTAGTGTCTCGAGCAAGAGGAAATTTGATGCTATTAACGCTAGCACTTGTCT	3660
Db	3601	GCTTAGTGTCTCGAGCAAGAGGAAATTTGATGCTATTAACGCTAGCACTTGTCT	3660
OY	3661	TACCTGTTTATTAACATTAAGCAATTTGGAGATTAATTAATGATTAACATTAAT	3720
Db	3661	TACCTGTTTATTAACATTAAGCAATTTGGAGATTAATTAATGATTAACATTAAT	3720
OY	3721	TTATGTCACAGTTTCTGAAACGATTTGATATCTAGATCTGGTATGACCCCATCT	3780
Db	3721	TTATGTCACAGTTTCTGAAACGATTTGATATCTAGATCTGGTATGACCCCATCT	3780
OY	3781	GGATTTGAAATCAAGTATATCTCCAGAGGCTTACCTTCAATTAAGCAATTAAC	3840
Db	3781	GGATTTGAAATCAAGTATATCTCCAGAGGCTTACCTTCAATTAAGCAATTAAC	3840
OY	3841	CATTAACCAAGATTTGATAGTATGATGATTAACAGATTAATTTGGACACTT	3900

3841 CATAAACGAGATTGATGCTATGATGATTAACGCAATTCATTGGGAGCTAT 3900  
3901 TTTCGAATTAAGTTGGATAGCTCAGTGGCTGTAATTAATGACGGAGGT 3960  
3901 TTTCTAATTAAGTTGGATAGCTCAGTGGCTGTAATTAATGACGGAGGT 3960  
3961 ACATGATGTCAGACATACAGGATGAGGTATGATATGATATGAGAGTGTACACCC 4020  
3961 ACATGATGTCAGACATACAGGATGAGGTATGATATGATATGAGAGTGTACACCC 4020  
4021 ACATTTGCTCACTAAATTCAAATATTCACAGCTGATATATGATGATTCATT 4080  
4021 ACATTTGCTCACTAAATTCAAATATTCACAGCTGATATATGATGATTCATT 4080  
4081 ATACCTGATGTTGTTATGATGATGATGATGATGATGATGATGATGATGATG 4140  
4081 ATACCTGATGTTGTTATGATGATGATGATGATGATGATGATGATGATGATG 4140  
4141 GTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4200  
4141 GTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4200  
4201 TGTCCAGAGAGCAATCCATCGGAGTGCAGGAGGCTCAGAGAGAGAGAGAGCT 4260  
4201 TGTCCAGAGAGCAATCCATCGGAGTGCAGGAGGCTCAGAGAGAGAGAGAGCT 4260  
4261 CACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4320  
4261 CACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4320  
4321 AAAATCTAAATTAAGTTGGATAGCTCAGTGGCTGTAATTAATGACGGAGGT 4380  
4321 AAAATCTAAATTAAGTTGGATAGCTCAGTGGCTGTAATTAATGACGGAGGT 4380  
4381 ATTAAGATGAGAGACTGCTGCTGTTAAGGCTGCTGATGATGATGATGATGAT 4440  
4381 ATTAAGATGAGAGACTGCTGCTGTTAAGGCTGCTGATGATGATGATGATGAT 4440  
4441 TGAATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 4500  
4441 TGAATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 4500  
4501 ATGCTATATATGCTTTGATGATGATGATGATGATGATGATGATGATGATGAT 4560  
4501 ATGCTATATATGCTTTGATGATGATGATGATGATGATGATGATGATGATGAT 4560  
4561 CTTTCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4620  
4561 CTTTCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4620  
4621 TATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4680  
4621 TATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4680  
4681 AAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4740  
4681 AAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4740  
4741 ATGAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4800  
4741 ATGAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4800  
4801 TCCCATGTCAG 4860  
4801 TCCCATGTCAG 4860  
4861 TTGCTACTTAACAG 4920  
4861 TTGCTACTTAACAG 4920  
4921 GCTCAG 4980  
4921 GCTCAG 4980

4981 GCTCAG 5040  
4981 GCTCAG 5040  
4981 GCTCAG 5040  
5041 GCTTGTGCTGCTGAG 5100  
5041 GCTTGTGCTGCTGAG 5100  
5101 CTGGAG 5160  
5101 CTGGAG 5160  
5161 ATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5220  
5161 ATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5220  
5221 CTTCTTTAAG 5280  
5221 CTTCTTTAAG 5280  
5281 AGCCAAATGTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5340  
5281 AGCCAAATGTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5340  
5341 CTGCTATGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5400  
5341 CTGCTATGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5400  
5401 GCAACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5460  
5401 GCAACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5460  
5461 CTCAATATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 5520  
5461 CTCAATATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 5520  
5521 TGGTGAATGAG 5580  
5521 TGGTGAATGAG 5580  
5581 ATGATATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5640  
5581 ATGATATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5640  
5641 GATTAATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 5700  
5641 GATTAATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 5700  
5701 TGAACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5760  
5701 TGAACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5760  
5761 CATTAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5820  
5761 CATTAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5820  
5821 TGAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5880  
5821 TGAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5880  
5881 AAGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5940  
5881 AAGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5940  
5941 CATTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6000  
5941 CATTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6000  
6001 CAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6060  
6001 CAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6060





```
QY 301 TATATGAAAAAGAGATCTCCAAACAAATTATATTTTGTATTAAGGAGAAATCGAAGCT 360
    |||
    |||
    |||
Db 301 TATATGAAAAAGAGAGATCTCCAAACAAATTATATTTTGTATTAAGGAGAAATCGAAGCT 360
    |||
    |||
    |||
QY 361 TGAAGGTCTTAATTGACAAACGAGACTACCTCCCTTCATATATGTATATATACATATA 420
    |||
    |||
    |||
Db 361 TGAAGGTCTTAATTGACAAACGAGACTACCTCCCTTCATATATGTATATATACATATA 420
    |||
    |||
    |||
QY 421 TACGATATATACGATATATACATATATACATATATATATATATATATATATATATAT 480
    |||
    |||
    |||
Db 421 TACGATATATACGATATATACATATATACATATATATATATATATATATATATATAT 480
    |||
    |||
    |||
QY 481 ATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 540
    |||
    |||
    |||
Db 481 ATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 540
    |||
    |||
    |||
QY 541 TATATATATGAGAGGATGAAGTACTATACGACTAGTCCATATATATATATATATATAT 600
    |||
    |||
    |||
Db 541 TATATATATGAGAGGATGAAGTACTATACGACTAGTCCATATATATATATATATATAT 600
    |||
    |||
    |||
QY 541 TATATATATGAGAGGATGAAGTACTATACGACTAGTCCATATATATATATATATATAT 600
    |||
    |||
    |||
Db 541 TATATATATGAGAGGATGAAGTACTATACGACTAGTCCATATATATATATATATATAT 600
    |||
    |||
    |||
QY 601 AGATATAGCATGACAAATATATATATATATATATATATATATATATATATATATATAT 660
    |||
    |||
    |||
Db 601 AGATATAGCATGACAAATATATATATATATATATATATATATATATATATATATATAT 660
    |||
    |||
    |||
QY 601 AGATATAGCATGACAAATATATATATATATATATATATATATATATATATATATATAT 660
    |||
    |||
    |||
Db 601 AGATATAGCATGACAAATATATATATATATATATATATATATATATATATATATATAT 660
    |||
    |||
    |||
QY 661 TAAAGACATATATATGATGATATATATATATATATATATATATATATATATATATAT 720
    |||
    |||
    |||
Db 661 TAAAGACATATATGATGATATATATATATATATATATATATATATATATATATATAT 720
    |||
    |||
    |||
QY 721 CTAAATTCACCTGATCCCTTTATGATATATATATATATATATATATATATATATATAT 780
    |||
    |||
    |||
Db 721 CTAAATTCACCTGATCCCTTTATGATATATATATATATATATATATATATATATATAT 780
    |||
    |||
    |||
QY 781 TATCTACTATACAGGACGACACACAACTACCTCACCAGTATGTTACAAACGATATA 840
    |||
    |||
    |||
Db 781 TATCTACTATACAGGACGACACACAACTACCTCACCAGTATGTTACAAACGATATA 840
    |||
    |||
    |||
QY 841 TTTTGCCCTTATGTTAATTTATATATATATATATATATATATATATATATATATATAT 900
    |||
    |||
    |||
Db 841 TTTTGCCCTTATGTTAATTTATATATATATATATATATATATATATATATATATATAT 900
    |||
    |||
    |||
QY 901 AATTTTCCAAACCCAAACCTCGGAGAAACCCCGTATGTTGCTGTAACAGAGACA 960
    |||
    |||
    |||
Db 901 AATTTTCCAAACCCAAACCTCGGAGAAACCCCGTATGTTGCTGTAACAGAGACA 960
    |||
    |||
    |||
QY 961 TCCGAGTCATTAATTTACACCAACAAACAAATTTAGACGACGAGCGCTTCCCAA 1020
    |||
    |||
    |||
Db 961 TCCGAGTCATTAATTTACACCAACAAACAAATTTAGACGACGAGCGCTTCCCAA 1020
    |||
    |||
    |||
QY 1021 TCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1080
    |||
    |||
    |||
Db 1021 TCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1080
    |||
    |||
    |||
QY 1081 AAACCCCTGGAATCCCCCCTCGGAGATCCATCGAGAGGATAGCCCCCGCATCCGCTCGGC 1140
    |||
    |||
    |||
Db 1081 AAACCCCTGGAATCCCCCCTCGGAGATCCATCGAGAGGATAGCCCCCGCATCCGCTCGGC 1140
    |||
    |||
    |||
QY 1141 GAGAGCGGATCCGATTCGCGATGAGAGGAGGATCTCCGATGAGAGGATCCGAGACC 1200
    |||
    |||
    |||
Db 1141 GAGAGCGGATCCGATTCGCGATGAGAGGAGGATCTCCGATGAGAGGATCCGAGACC 1200
    |||
    |||
    |||
QY 1201 ATTTTGGGTCCTCCCTCCGCGCGAGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 1260
    |||
    |||
    |||
Db 1201 ATTTTGGGTCCTCCCTCCGCGCGAGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 1260
    |||
    |||
    |||
QY 1261 CGGCGGCTTCGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
    |||
    |||
    |||
Db 1261 CGGCGGCTTCGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
    |||
    |||
    |||
QY 1321 GAACCGGTGCTCCGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
    |||
    |||
    |||
Db 1321 GAACCGGTGCTCCGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
    |||
    |||
    |||
QY 1381 CCCCGTCCGAAACCTTAGCCCGAGGCGCAAGCGGAGAGGATCAGGCGCGCAGAGGAGGT 1440
    |||
    |||
    |||
```



Dd	2461	TCATTTTGGTACATGACTTTAAATAGACATAGGACCAATGCATGATGTTGGTTATTCAT	2520
Oy	2521	TAATCATTTATGATAGTTCCTGTCATTTATCATATGCTATGCTACATGATCATTTTCATCTCTCAGG	2580
Dd	2522	AAATCATTTATGATAGTTCCTGTCATTTATCATATGCTATGCTACATGATCATTTTCATCTCTCAGG	2580
Oy	2581	CCTCTGGACATGTTCCACTGAGCGTCCCTGGACCTGGTTCATCCTTGGTAAATGACATG	2640
Dd	2581	CCTCTGGACATGTTCCACTGAGCGTCCCTGGACCTGGTTCATCCTTGGTAAATGACATG	2640
Oy	2641	TTTTACACATAGGCGGCTCCTATTCATCATGAGAGTACTATCTATCTATCTGGTATTCAT	2700
Dd	2641	TTTTACACATAGGCGGCTCCTATTCATCATGAGAGTACTATCTATCTATCTGGTATTCAT	2700
Oy	2701	AGATTGTTATGAAACTACCCAAATATCTCCACATTCGATGGGATTTAAATTTAGTTTC	2760
Dd	2701	AGATTGTTATGAAACTACCCAAATATCTCCACATTCGATGGGATTTAAATTTAGTTTC	2760
Oy	2761	TTTTGAATATAGATATAGATGTTATTTGGCGACAGTCATCAATTAATTTAGTTCTGAACATATGA	2820
Dd	2761	TTTTGAATATAGATATAGATGTTATTTGGCGACAGTCATCAATTAATTTAGTTCTGAACATATGA	2820
Oy	2821	TTAATATAGTTCCGCAATTTGTTATGATTCCTTGAACATTAATTAATTTGATTCCTTAAGTNG	2880
Dd	2821	TTAATATAGTTCCGCAATTTGTTATGATTCCTTGAACATTAATTAATTTGATTCCTTAAGTNG	2880
Oy	2881	ATAAGGTTATGTTGTTTGGAGATCCCTTAATCATCTATTCATGCTGAGCTCCAGCGTGA	2940
Dd	2881	ATAAGGTTATGTTGTTTGGAGATCCCTTAATCATCTATTCATGCTGAGCTCCAGCGTGA	2940
Oy	2941	GGAGCATTTGTTGGCTTGCATACATGAAATGAAGACCTCCTGTTCTGATGCTGCACAG	3000
Dd	2941	GGAGCATTTGTTGGCTTGCATACATGAAATGAAGACCTCCTGTTCTGATGCTGCACAG	3000
Oy	3001	AAATATATTTTATTAATGCACTGTAATCCTTAAGATCTTGCATTAAGATGGGCACTTNG	3060
Dd	3001	AAATATATTTTATTAATGCACTGTAATCCTTAAGATCTTGCATTAAGATGGGCACTTNG	3060
Oy	3061	CTTTGGAATGATAGTACTATTAATAGCTGTGTTATTCATGATTTGGCTGTGATGAAT	3120
Dd	3061	CTTTGGAATGATAGTACTATTAATAGCTGTGTTATTCATGATTTGGCTGTGATGAAT	3120
Oy	3121	GTGCACAAAACAGGAATGCTACTTTTGTACTTCTGATATTTCAATGCTGGCTTACTTATG	3180
Dd	3121	GTGCACAAAACAGGAATGCTACTTTTGTACTTCTGATATTTCAATGCTGGCTTACTTATG	3180
Oy	3181	CTCTTTTGGAAATGAGGACATATACAGGACATGACACCGATGCCAAACATGCTAAAT	3240
Dd	3181	CTCTTTTGGAAATGAGGACATATACAGGACATGACACCGATGCCAAACATGCTAAAT	3240
Oy	3241	GGCCCAAGTGGGGATCGGGCTCACAGTTGGTACAGAAATGTTGATGTCCTGTTAAAGCAG	3300
Dd	3241	GGCCCAAGTGGGGATCGGGCTCACAGTTGGTACAGAAATGTTGATGTCCTGTTAAAGCAG	3300
Oy	3301	CCACACAGCTTCCTTCATCAAGGAGAGAGTACAGATGAGATGATTAAGAGAGAAAGCTGAG	3360
Dd	3301	CCACACAGCTTCCTTCATCAAGGAGAGAGTACAGATGAGATGATTAAGAGAGAAAGCTGAG	3360
Oy	3361	ACCACTGGAAATCGAAGACCTGCTATCAAAAGATTCAACGAAAGGTGATATTCATTTGCT	3420
Dd	3361	ACCACTGGAAATCGAAGACCTGCTATCAAAAGATTCAACGAAAGGTGATATTCATTTGCT	3420
Oy	3421	TGCTTGAATATAGATTTCTGTACATTAATTAACCTACCTGCTATCAACGATGTGTGTCTA	3480
Dd	3421	TGCTTGAATATAGATTTCTGTACATTAATTAACCTACCTGCTATCAACGATGTGTGTCTA	3480
Oy	3481	TTTTTACCACTAGCCCTTTAGTTGATTTCCACTTTTCAACCGGATAGCCTTTCAAGTTTCT	3540
Dd	3481	TTTTTACCACTAGCCCTTTAGTTGATTTCCACTTTTCAACCGGATAGCCTTTCAAGTTTCT	3540
Oy	3541	ATTGATATGCAATATATGATCTTTTACCTACCAATATATGTTCTCTGTGTGCATATCACT	3600
Dd	3541	ATTGATATGCAATATATGATCTTTTACCTACCAATATATGTTCTCTGTGTGCATATCACT	3600

QY	3601	GCTTAGTGTCTCAGACAGAGAGAAATTTGTATAGGCTATATACAGTGTACACTTGTGCTTC	3660
Db	3601	GCTTAGTGTCTCAGACAGAGAGAAATTTGTATAGGCTATATACAGTGTACACTTGTGCTTC	3660
QY	3661	TACTGTATTATGACAATAAGCAATTGGAGATGAATTAATCCGAGTTGCATCATATCTTC	3720
Db	3661	TACTGTATTATGACAATAAGCAATTGGAGATGAATTAATCCGAGTTGCATCATATCTTC	3720
QY	3721	TTATGTCAAGTTTGTGAACCGATGTATCTAGTATCGTGTGATGACCCCCACTT	3780
Db	3721	TTATGTCAAGTTTGTGAACCGATGTATCTAGTATCGTGTGATGACCCCCACTT	3780
QY	3781	GGATTTGGAAATAAAGTATATCTCCCTAGAGAGCTTTCACCTTCATTAAGCAATTAACC	3840
Db	3781	GGATTTGGAAATAAAGTATATCTCCCTAGAGAGCTTTCACCTTCATTAAGCAATTAACC	3840
QY	3841	CAATTAACACAGGATTTGATAGCTATAGACTATAGATTAACAGAAATTCATTTGGACACTAT	3900
Db	3841	CAATTAACACAGGATTTGATAGCTATAGACTATAGATTAACAGAAATTCATTTGGACACTAT	3900
QY	3901	TTTTCATTTAAGTTTGGTATATATCTCAGTTGGCTGTAAATATATGTCAAGCTAGGCT	3960
Db	3901	TTTTCATTTAAGTTTGGTATATATCTCAGTTGGCTGTAAATATATGTCAAGCTAGGCT	3960
QY	3961	ACANTATATGTACAGCAATACAAAGTATGGGTCAGTTATGATATAGCAAGTGTGACACCCC	4020
Db	3961	ACANTATATGTACAGCAATACAAAGTATGGGTCAGTTATGATATAGCAAGTGTGACACCCC	4020
QY	4021	ACATTTGCTCACTAAAATATAATATTCACCTCCAGTATATGTGTGATGAT	4080
Db	4021	ACATTTGCTCACTAAAATATAATATTCACCTCCAGTATATGTGTGATGAT	4080
QY	4081	ATACCTGTATTTGTTATATATGTACTGTGTGACAGACAATATATAGGCTGTCTTTGG	4140
Db	4081	ATACCTGTATTTGTTATATATGTACTGTGTGACAGACAATATATAGGCTGTCTTTGG	4140
QY	4141	GTGATTTTGTATAGAGATTTGAGCAAGCACTTCTCGATATATATCTAGTTTGTGTAC	4200
Db	4141	GTGATTTTGTATAGAGATTTGAGCAAGCACTTCTCGATATATATCTAGTTTGTGTAC	4200
QY	4201	TTGTTCAGAGAAAGCAATTCACATCGGAGTCAGCCAGAGCGCTCAAGAAAGCAAGAGCAAGT	4260
Db	4201	TTGTTCAGAGAAAGCAATTCACATCGGAGTCAGCCAGAGCGCTCAAGAAAGCAAGAGCAAGT	4260
QY	4261	CACCTGAATGAAGTCAGAGGACAGAGTGTGATTCACATATCTATTTTTCATTAAGCAT	4320
Db	4261	CACCTGAATGAAGTCAGAGGACAGAGTGTGATTCACATATCTATTTTTCATTAAGCAT	4320
QY	4321	AAAAACCTAATATACGTGCTACTGACTGACTTCAGTTATGTGATTCAGTATACGATACGA	4380
Db	4321	AAAAACCTAATATACGTGCTACTGACTGACTTCAGTTATGTGATTCAGTATACGATACGA	4380
QY	4381	ATTATAGAGTCAGAACTCTCGCTTATAGGCGCTCTCTGATTTATACAGAGATACAA	4440
Db	4381	ATTATAGAGTCAGAACTCTCGCTTATAGGCGCTCTCTGATTTATACAGAGATACAA	4440
QY	4441	TCATGCTCTCTGTGACATATAGAGTCTTAAAGAGATTTGAGACCTTGTAGACACCAAGCT	4500
Db	4441	TCATGCTCTCTGTGACATATAGAGTCTTAAAGAGATTTGAGACCTTGTAGACACCAAGCT	4500
QY	4501	ATGCTATATATGCTTTTTCATATATGATCCCATAGATATGATATTTGGCTGTCTTCAA	4560
Db	4501	ATGCTATATATGCTTTTTCATATATGATCCCATAGATATGATATTTGGCTGTCTTCAA	4560
QY	4561	CTTTCACAGTACTTGTGACCTGTATTTAGAGAAATATCCGCTTACCATTAATATCTC	4620
Db	4561	CTTTCACAGTACTTGTGACCTGTATTTAGAGAAATATCCGCTTACCATTAATATCTC	4620
QY	4621	TTTAAATACCATTTGGCCAGTCACAAATAGATTTATTAATAGGCTCAATGTGACATTG	4680
Db	4621	TTTAAATACCATTTGGCCAGTCACAAATAGATTTATTAATAGGCTCAATGTGACATTG	4680

QY	4681	AAATGCTATCAAATATTCGAAAGTGAAGATGGCAGAGACTCGGTGAAGCGGGTGAACGGC	4740
Db	4681	AAATGCTATCAAATATTCGAAAGTGAAGATGGCAGAGACTCGGTGAAGCGGGTGAAGCGC	4740
QY	4741	ATGAACGGGTGTTGTTCCCGCGCTTCGATATATCAATCCACATGCCATTCACACAC	4800
Db	4741	ATGAACGGGTGTTTCCCGCGCTTCCGACCTTCGATATATCAATCCACATGCCATTCACACAC	4800
QY	4801	TTCCCATCTGAAGCAACGTCAAGCGCTGCTTCCCATCTGAAGTGAACCCGACATATAC	4860
Db	4801	TTCCCATCTGAAGCAACGTCAAGCGCTGCTTCCCATCTGAAGTGAACCCGACATATAC	4860
QY	4861	TTTCGCTACTAACAAGCACAATCGAGAGTAAACAATATACATACCCGACATACCTTTTCG	4920
Db	4861	TTTCGCTACTAACAAGCACAATCGAGAGTAAACAATATACATACCCGACATACCTTTTCG	4920
QY	4921	GCTCAGAGAGACGAGACATCTGCATAATGGCGCTCGCTCGCGGCAAGATTTGGCCGGCA	4980
Db	4921	GCTCAGAGAGACGAGACATCTGCATAATGGCGCTCTGCGCGGCAAGATTTGGCCGGCA	4980
QY	4981	GCGTCGCTGAGCGGGTGGCGAGCTCGGTGAACAATCCACAAAGAGATGTCGGTGGCGC	5040
Db	4981	GCGTCGCTGAGCGGGTGGCGAGCTCGGTGAACAATCCACAAAGAGATGTCGGTGGCGC	5040
QY	5041	GCTTCGCTGGGTGCAAGCTCTGAGACCGAAACCCAGACGCTTCGCTTCAAGACA	5100
Db	5041	GCTTCGCTGGGTGCAAGCTCTGAGACCGAAACCCAGACGCTTCGCTTCAAGACA	5100
QY	5101	CTGGCAGACAGAAATATATATTTGGACAGCGGTAGACATCTTACTGCTGAGAGTTGATG	5160
Db	5101	CTGGCAGACAGAAATATATATTTGGACAGCGGTAGACATCTTACTGCTGAGAGTTGATG	5160
QY	5161	ATTGTGGTGTGCTGCTGCTGATCTGATTTGCAATTTCTTGAGAGACACTCTTGATACCT	5220
Db	5161	ATTGTGGTGTGCTGCTGCTGATCTGATTTGCAATTTCTTGAGAGACACTCTTGATACCT	5220
QY	5221	CTTCTCTAAGGACCTTAGTACCCCTTGGGATCTCATCTGCTCTTAATACTATACATT	5280
Db	5221	CTTCTCTAAGGACCTTAGTACCCCTTGGGATCTCATCTGCTCTTAATACTATACATT	5280
QY	5281	AGCAAAATGTTCAATTTGGTGTGAGCGCTGCTCTTAATGGAAGACATGTTTCAAGCAG	5340
Db	5281	AGCAAAATGTTCAATTTGGTGTGAGCGCTGCTCTTAATGGAAGACATGTTTCAAGCAG	5340
QY	5341	CTGCTATGCTATCATTCAATAATATTTGATGATGATGCTCTTCTGCTTTGCTTTAA	5400
Db	5341	CTGCTATGCTATCATTCAATAATATTTGATGATGATGCTCTTCTGCTTTGCTTTAA	5400
QY	5401	GCAACCAAGATTAATATATCAACCTTTTGAGCGTCAATTGAGTGCAGATAG	5460
Db	5401	GCAACCAAGATTAATATATCAACCTTTTGAGCGTCAATTGAGTGCAGATAG	5460
QY	5461	CTCAATATCTCAGGTGTCATTGAGTTTAAAGTGGAACGATTAACAAACGTCAGGCTA	5520
Db	5461	CTCAATATCTCAGGTGTCATTGAGTTTAAAGTGGAACGATTAACAAACGTCAGGCTA	5520
QY	5521	TGCGAATGAGAGGAGAGGTATACATCCCTAATATATCTCAATTTTCAATATTAATTTAGTG	5580
Db	5521	TGCGAATGAGAGGAGAGGTATACATCCCTAATATATCTCAATTTTCAATATTAATTTAGTG	5580
QY	5581	ATGCAATTTTACCAACCAATCCCATCAAGTCAATCAATCAACAGAGGTATATTAATG	5640
Db	5581	ATGCAATTTTACCAACCAATCCCATCAAGTCAATCAATCAACAGAGGTATATTAATG	5640
QY	5641	GTAATATTTTCGTTTGAAGAAAAAAGAGAACCTTATATTAAGTTTACCGGTGGGGTG	5700
Db	5641	GTAATATTTTCGTTTGAAGAAAAAAGAGAACCTTATATTAAGTTTACCGGTGGGGTG	5700
QY	5701	TGAACATTAATCAATGAAGAGATGCAATCCGCTAGGAGGAGCTACTGTGAACAAATAG	5760
Db	5701	TGAACATTAATCAATGAAGAGATGCAATCCGCTAGGAGGAGCTACTGTGTGAACAAATAG	5760
QY	5761	CATTAATCTCGGTATCCACCAACAAACGCTTGGCGACGGCGTCAAAATGGCAGGACT	5820

Db	5761	CATAAACTCCGATACCAACACAAACAGCTGGGACCGCTCTAAAGGCGAGGACT	5820		
Qy	5821	TCATCGCTTCGGGGCAAGAAAGAAATCAAGTGAATACATTGGCAGGGAAACCCAAAG	5880		
Db	5821	TCATCGCTTCGGGGCAAGAAAGAAATCAAGTGAATACATTGGCAGGGAAACCCAAAG	5880		
Qy	5881	AAGGCATCCAAATCCATTCACATCCACTCCAAACGGGCGATGGAAACAAGACGATGATTCACAG	5940		
Db	5881	AAGGCATCCAAATCCATTCACATCCACTCCAAACGGGCGATGGAAACAAGACGATGATTCACAG	5940		
Qy	5941	CTATCTCTCTCTCTCAAGTTTGAATCTTGTACGTGCTTCAGGGAAAAAGAGAT	6000		
Db	5941	CTATCTCTCTCTCTCAAGTTTGAATCTTGTACGTGCTTCAGGGAAAAAGAGAT	6000		
Qy	6001	CAGATTAAGTGTGATTCGCGGGCGGCTGAGTTCTTGGGAGATCTTGTGTGGAGTGGCA	6060		
Db	6001	CAGATTAAGTGTGATTCGCGGGCGGCTGAGTTCTTGGGAGATCTTGTGTGGAGTGGCA	6060		
Qy	6061	GGAGGAGCAATCGGCTGGCCGCTTCTGCTACCGAAACATCGCCAGTAAAGGCCAA	6120		
Db	6061	GGAGGAGCAATCGGCTGGCCGCTTCTGCTACCGAAACATCGCCAGTAAAGGCCAA	6120		
Qy	6121	AAGCAATTAATTAACGCATAGGGGAGATCGCCCATCTGCATTAACAATTCGATGACGGAATCG	6180		
Db	6121	AAGCAATTAATTAACGCATAGGGGAGATCGCCCATCTGCATTAACAATTCGATGACGGAATCG	6180		
Qy	6181	ATTAAATACAAAGATGACATGTGATTAATTACGCTGCGAAGCTT	6227		
Db	6181	ATTAAATACAAAGATGACATGTGATTAATTACGCTGCGAAGCTT	6227		
RESULT 3					
ID	AAD45326	AAD45326 standard; DNA: 3902 BP.			
AC	AAD45326;				
XX					
DT	27-DEC-2002	(first entry)			
XX					
DE	Rice endosperm bZip (Rb) transcription factor DNA.				
XX					
KW	Feed improvement; plant-derived feed; antibiotic; additive; flour;				
KW	anti-microbial; poultry; alpha-1-antitrypsin; rice endosperm 2 bZip;				
KW	Rb; malt; transcription factor; ds.				
XX					
OS	Oryza sativa.				
XX					
PN	W0200263975-A2.				
XX					
PD	22-AUG-2002.				
XX					
PF	14-FEB-2002; 2002W0-US04919.				
XX					
PR	14-FEB-2001; 2001US-269188P.				
XX					
PR	02-MAY-2001; 2001US-0847232.				
XX					
PA	(VENT-) VENTRIA BIOSCIENCE.				
XX					
PI	Huang N, Rodriguez RL, Hagie FE;				
XX					
DR	WPI; 2002-682708/73.				
XX					
XX					
PT	Improved feed for production animals, comprising plant-derived feed				
PT	ingredients, and seed composition containing flour, extract, or malt				
PT	from mature monoot seeds and heterologous seed-produced anti-microbial				
PT	proteins -				
XX					
PS	Disclosure: Page 159-160; 175pp; English.				
XX					
CC	The invention relates to improved feed for production animals,				
CC	comprising one or more plant-derived feed ingredients, substantially				
CC	unsupplemented with small-molecule antibiotics and as an additive a				

CC seed composition containing a flour, extract or malt obtained from  
CC mature monocot seeds and one or more heterologous seed-produced anti-  
CC microbial proteins in substantially unpurified form. The invention  
CC is useful as a feed for production animals such as poultry and hoofed  
CC farm animals. The present sequence is rice endosperm 23ip (reb)  
CC transcription factor DNA. This sequence is used in the invention.  
XX

Sequence 3902 BP; 975 A; 824 C; 937 G; 1166 T; 0 other:

Query Match 62.7%; Score 3902; DB 24; Length 3902;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 3902; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1163 ATGAGACGGGTGTTCTCCGTGGAGAGATCTCCAGACCATTTGGGTCCCGCTCCGCG 1222

DB 1 ATGAGACGGGTGTTCTCCGTGGAGAGATCTCCAGACCATTTGGGTCCCGCTCCGCG 60

QY 1223 CCGCAGTGGCGCGCGCGCCAGACAGAGGGCGCGCGCGCTTGGGAGGTGT 1282

DB 61 CCGCAGTGGCGCGCGCGCCAGACAGAGGGCGCGCGCGCTTGGGAGGTGT 120

QY 1283 GGTGGTGAAGCGGGGCGCGCGCGCGAGACGATGAACCGGTGCCCGCGAGTGG 1342

DB 121 GGTGGTGAAGCGGGGCGCGCGCGCGAGACGATGAACCGGTGCCCGCGAGTGG 180

QY 1343 TACTTCAGAAATTTCTGAGAGAGCGGTCTCGATAGCCCGCTCCGACCCCTAGCCCG 1402

DB 181 TACTTCAGAAATTTCTGAGAGAGCGGTCTCGATAGCCCGCTCCGACCCCTAGCCCG 240

QY 1403 AGGGCGGAAGCGGGAGGATCAGGGGCGAGAGGGGTGTGGCGCGGTCTGAGTGAAGAG 1462

DB 241 AGGGCGGAAGCGGGAGGATCAGGGGCGAGAGGGGTGTGGCGCGGTCTGAGTGAAGAG 300

QY 1463 CCGCAGCTCTCGCGCGCGCGAGACAGCGCGGTGTGGAGCCCTGGAGTACAACGCG 1522

DB 301 CCGCAGCTCTCGCGCGCGAGACAGCGCGGTGTGGAGCCCTGGAGTACAACGCG 360

QY 1523 ATGCTGAAGCAGAACTGAGAGAAAGACTCCGCGCGGTCCGCACTGGAGGCTACAGCA 1582

DB 361 ATGCTGAAGCAGAACTGAGAGAAAGACTCCGCGCGGTCCGCACTGGAGGCTACAGCA 420

QY 1583 TTCTCCCCCCTCTAGACTCGAGACTTACTGAGATCGGCAATGCTACTCTCTCTTGGC 1642

DB 421 TTCTCCCCCCTCTAGACTCGAGACTTACTGAGATCGGCAATGCTACTCTCTCTTGGC 480

QY 1643 ATCGAATGTTATATGATATTAGATCGGCAATTTCTATTAACCAATGCGCTCCATGCTCT 1702

DB 481 ATCGAATGTTATATGATATTAGATCGGCAATTTCTATTAACCAATGCGCTCCATGCTCT 540

QY 1703 TCGAATGCGCTGTGTAGTGTGCGGTGTGGTGTGACTCATATGATGTAGAGGTTGTGCG 1762

DB 541 TCGAATGCGCTGTGTAGTGTGCGGTGTGGTGTGACTCATATGATGTAGAGGTTGTGCG 600

QY 1763 TATGTACAAGCGAAGCTTCATAGACCTCGATATGAGATTTGGATATGATGATCAACCTG 1822

DB 601 TATGTACAAGCGAAGCTTCATAGACCTCGATATGAGATTTGGATATGATGATCAACCTG 660

QY 1823 CGAATTTGGCGATATATCACTATTTCTTAACTCGAGAGACGTGTTGTTTCA 1882

DB 661 CGAATTTGGCGATATATCACTATTTCTTAACTCGAGAGACGTGTTGTTTCA 720

QY 1883 ATGCAATATTTTGTATGGGCTGTCTTAACTGTATGCTGCTTTTAAATTTGGCAATAT 1942

DB 721 ATGCAATATTTTGTATGGGCTGTCTTAACTGTATGCTGCTTTTAAATTTGGCAATAT 780

QY 1943 GTGACTTTATGCAAGTATTTGATTTGGCGGATCAAGAAACAAAAGTTGGGGGATTCAA 2002

DB 781 GTGACTTTATGCAAGTATTTGATTTGGCGGATCAAGAAACAAAAGTTGGGGGATTCAA 840

QY 2003 CATACGAGTACACTGCGCATTAACATATCTGATTTAACTATGCTTAAAGATCAT 2062

DB 841 CATACGAGTACACTGCGCATTAACATATCTGATTTAACTATGCTTAAAGATCAT 900

QY 2063 TTAGAGACCTTTAGACACTCTTATCTTATCAACATGATGAAATAATGAAAGGGGAGCT 2122

DB 901 TTAGAGACCTTTAGACACTCTTATCTTATCAACATGATGAAATAATGAAAGGGGAGCT 960

QY 2123 CAGGGGGATTCATATGGGTCCGATGGGTGAGAGGGGGAATGATGTCCTCCCGACCCACG 2182

DB 961 CAGGGGGATTCATATGGGTCCGATGGGTGAGAGGGGGAATGATGTCCTCCCGACCCACG 1020

QY 2183 TTGAATCCGCCCTGCGCATGCTATAGCTGTCAACACCATTTCTAGGTGCTGTGTAG 2242

DB 1021 TTGAATCCGCCCTGCGCATGCTATAGCTGTCAACACCATTTCTAGGTGCTGTGTAG 1080

QY 2243 TTGGGTATGTACGCTTAATTTGCTTTCTATGTGCTCATGATTTCTTAAGAACGA 2302

DB 1081 TTGGGTATGTACGCTTAATTTGCTTTCTATGTGCTCATGATTTCTTAAGAACGA 1140

QY 2303 AAATAGCCTATTTATGTGCTCCAGAAATTTGATATCCCGGCCCTCATTTGCTGAATT 2362

DB 1141 AAATAGCCTATTTATGTGCTCCAGAAATTTGATATCCCGGCCCTCATTTGCTGAATT 1200

QY 2363 AGCCTATTTGTTGTTGCCCTTCAGTTTTCACAGCTTATGTTGTCATATGTGCT 2422

DB 1201 AGCCTATTTGTTGTTGCCCTTCAGTTTTCACAGCTTATGTTGTCATATGTGCT 1260

QY 2423 ATGCTGCTTTTGTGCTCCATATTTATTTATTTGCAATTCATTTTGTACATGACTTAA 2482

DB 1261 ATGCTGCTTTTGTGCTCCATATTTATTTATTTGCAATTCATTTTGTACATGACTTAA 1320

QY 2483 ATGACACTAAGCAATCATGACTGATTTGTTATCATATATATATTTATGATTTCTGTTG 2542

DB 1321 ATGACACTAAGCAATCATGACTGATTTGTTATCATATATATATTTATGATTTCTGTTG 1380

QY 2543 ATTTATCATATGCTATGCTATGCTATTTTCAATCTTCAAGGCTGTGACAGATTTCCACTGTA 2602

DB 1381 ATTTATCATATGCTATGCTATGCTATTTTCAATCTTCAAGGCTGTGACAGATTTCCACTGTA 1440

QY 2603 GCGTCCGAGCGCTGCTATCTTCTGTAAGCAAGATGTTTCAACATAGGCGCTCTAA 2662

DB 1441 GCGTCCGAGCGCTGCTATCTTCTGTAAGCAAGATGTTTCAACATAGGCGCTCTAA 1500

QY 2663 TTCCATATGAGGACTTATCTTATCTGTTACATTTTCAATTTGATTAACATACCA 2722

DB 1501 TTCCATATGAGGACTTATCTTATCTGTTACATTTTCAATTTGATTAACATACCA 1560

QY 2723 ATATCTGACACATTTGATGAGGATTAATTTATTTCTTTGAAATAGAGATGATGT 2782

DB 1561 ATATCTGACACATTTGATGAGGATTAATTTATTTCTTTGAAATAGAGATGATGT 1620

QY 2783 ATTTGCTGACGTCATCAATATGTTCTGAACCTTGAATTAATAGTTCCGATTTGTTA 2842

DB 1621 ATTTGCTGACGTCATCAATATGTTCTGAACCTTGAATTAATAGTTCCGATTTGTTA 1680

QY 2843 GTGATTTCTTGAACATTAATATTTATGCTTAAATAGATAGGTTATGTTTGTGGAG 2902

DB 1681 GTGATTTCTTGAACATTAATATTTATGCTTAAATAGATAGGTTATGTTTGTGGAG 1740

QY 2903 TTCCCTTAATCATTTGATTCGAGCTCCGACGCGAGAGGACTTGTGTGCGCTGA 2962

DB 1741 TTCCCTTAATCATTTGATTCGAGCTCCGACGCGAGAGGACTTGTGTGCGCTGA 1800

QY 2963 CCATGAATGAAGACCTTCTGTTGATGCTGACCAAGAAACATATTTTGAATATCA 3022

DB 1801 CCATGAATGAAGACCTTCTGTTGATGCTGACCAAGAAACATATTTTGAATATCA 1860

QY 3023 CCTTAAATCCTTAGATCTTGAAGATGGGACCTTACCTTTGATTTAGATGATGACTTA 3082

DB 1861 CCTTAAATCCTTAGATCTTGAAGATGGGACCTTACCTTTGATTTAGATGATGACTTA 1920

QY 3083 AATAGCTGTGTTATGCTATGTTTGTGCTGATGTAATGTCGACAAACAGGAATGCTAC 3142

DB 1921 AATAGCTGTGTTATGCTATGTTTGTGCTGATGTAATGTCGACAAACAGGAATGCTAC 1980

QY 3143 TTTTGAATTCGATATTTCAATGCTGCTTTACTTATGCTGTTTGAACATGGGACCA 3202

Db	1981	TTTTGACCTCGAATTATTCATGAGCCGCTTACCTATATGCTCGTTTGGAACATAGGACACA	2040
QY	3203	TATCAGGCAATGCTACTCCAGTTCAAAAACATGCTAATGTCGACCAAGTGGGGATCGGGCT	3262
Db	2041	TATCAGGCAATGCTACTCCAGTTCAAAAACATGCTAATGTCGACCAAGTGGGGATCGGGCT	2100
QY	3263	CACAGTTGGACAGAAATGTGATGCTGTAAAGAGAGCCACAGCTCTTCATCAAGG	3322
Db	2101	CACAGTTGGACAGAAATGTGATGCTGTAAAGAGAGCCACAGCTCTTCATCAAGG	2160
QY	3323	AGCAGTCAGATGATGATGACATGGAAGGAGAAAGCTGAGACATGGAATGCAAGACTCG	3382
Db	2161	AGCAGTCAGATGATGATGACATGGAAGGAGAAAGCTGAGACATGGAATGCAAGACTCG	2220
QY	3383	CGAATCGAAATATTAACAAGAAAGGATCATCTCTGCTCTTAAATATAGATCTCGTA	3442
Db	2221	CTGATCCAAATTTACAGAGAAAGGATCATCTCTGCTCTTAAATATAGATCTCGTA	2280
QY	3443	CATATATTAACCTACCTGCTCATGATCATGTCCTCTATTTTCACCTTACCTTTACGT	3502
Db	2281	CATATATTAACCTACCTGCTCATGATCATGTCCTCTATTTTCACCTTACCTTTACGT	2340
QY	3503	TGGATTTCCACTTTACCGGGATGACCTTTAGTTTCTATTCATGCCATATATGATCTT	3562
Db	2341	TGGATTTCCACTTTACCGGGATGACCTTTAGTTTCTATTCATGCCATATATGATCTT	2400
QY	3563	TTACCTACCATATTAATTCCTGTGTGCCATATCATAGTCTTAGTCTGACACAAGAA	3622
Db	2401	TTACCTACCATATTAATTCCTGTGTGCCATATCATAGTCTTAGTCTGACACAAGAA	2460
QY	3623	GGATTTTGTATGGCTATTTACAGTGAACCTTTGCTCTACTGTTTATTTGACATAGCA	3682
Db	2461	GGATTTTGTATGGCTATTTACAGTGAACCTTTGCTCTACTGTTTATTTGACATAGCA	2520
QY	3683	ATTGGGATTAATTAATTCGTAGTTCATCATCATATATTCCTATGTCACAAATTCCTAAAC	3742
Db	2521	ATTGGGATTAATTAATTCGTAGTTCATCATCATATATTCCTATGTCACAAATTCCTAAAC	2580
QY	3743	CGATTTGATCTAGTATCTGTTGTGATGACACCCCATCTTGATTTGCCAAATCAAAAGTTATA	3802
Db	2581	CGATTTGATCTAGTATCTGTTGTGATGACACCCCATCTTGATTTGCCAAATCAAAAGTTATA	2640
QY	3803	CMCCCTGAGAGGCTTAACCTTCATATAAGCAATTAACCCATAAACCAAGATTTATAG	3862
Db	2641	CTCCCTGAGAGGCTTAACCTTCATATAAGCAATTAACCCATAAACCAAGATTTATAG	2700
QY	3863	CTATTGACATGATTAACAGAAATTCATTGGCAGCAATTTCTCAATTTAGTTTGAT	3922
Db	2701	CTATTGACATGATTAACAGAAATTCATTGGCAGCAATTTCTCAATTTAGTTTGAT	2760
QY	3923	TAGTCCTAGTTGGCTGTAAATATATATATGTCACGGATGACATATGATGCGACATACAG	3982
Db	2761	TAGTCCTAGTTGGCTGTAAATATATATATGTCACGGATGACATATGATGCGACATACAG	2820
QY	3983	GTATGGGAGGATTAATATATGAGACAGGTGTGAACCCCATTTGCTCAGCTAAATCAAA	4042
Db	2821	GTATGGGAGGATTAATATATGAGACAGGTGTGAACCCCATTTGCTCAGCTAAATCAAA	2880
QY	4043	ATATTCAAACGTCAGTGATGATAGTGGATTCGATTAACCTTGATTTGTTATTATG	4102
Db	2881	ATATTCAAACGTCAGTGATGATAGTGGATTCGATTAACCTTGATTTGTTATTATG	2940
QY	4103	TTACCTTGCGTAAACAATAATATAGCGCTTTTGGGTGATTTTGTATGAAGATGTG	4162
Db	2941	TTACCTTGCGTAAACAATAATATAGCGCTTTTGGGTGATTTTGTATGAAGATGTG	3000
QY	4163	AGCAAGACATTCGATATAATAGTACGTTTGTGACGCTTCACAGAAACAATCCATAC	4222
Db	3001	AGCAAGACATTCGATATAATAGTACGTTTGTGACGCTTCACAGAAACAATCCATAC	3060
QY	4223	GGGAGTCAGCCAGGCGCTTACAGACGAAAGGCACTACTTGATATGAGCTGGAGCAC	4282

Db	3061	56GAGTCCAGCCAGCGCTCCAGAGGCAAGAAAGGCAAGCTCCCTTAATGAGCTGAGGAC	3120
OY	4283	AGGTGTATAGTTCACATAGTTATTTCGTAAAGACATPAAAATCCATAATACGGCTAC	4342
Db	3121	AGGTGTATAGTTCACATAGTTATTTCGTAAAGACATPAAAATCCATAATACGGCTAC	3180
OY	4343	TGACCTCAGTATGAGTTTACTGTGTACAGTATGCGAATTAAGTCGAGAACCTCCG	4402
Db	3181	TGACCTCAGTATGAGTTTACTGTGTACAGTATGCGAATTAAGTCGAGAACCTCCG	3240
OY	4403	CTGTAAAGGCGTCTTCTGATGTATTAACAGAACTACAAATGATGCTGTGTACATAGA	4462
Db	3241	CTGTAAAGGCGTCTTCTGATGTATTAACAGAAATPAATAATGATGCTGTGTACATAGA	3300
OY	4463	GTGTCTAAAGCAATGTGAGACCTTGAGGCAAGAGATGCTATATATGCTTTGGAA	4522
Db	3301	GTGTCTAAAGCAATGTGAGACCTTGAGGCAAGAGATGCTATATATGCTTTGGAA	3360
OY	4523	TATGATCCCATGATGATGCTACTTGGCTGTTCCTAACTTCACCTGACCTGTGTAC	4582
Db	3361	TATGATCCCATGATGATGCTACTTGGCTGTTCCTAACTTCACCTGACCTGTGTAC	3420
OY	4583	CTGTATTATTAATAAATTAATCCCGCTACCTTTACTGTATTAATACACTTTGGCCAGT	4642
Db	3421	CTGTATTATTAATAAATAAATCCCGCTACCTTTACTGTATTAATACACTTTGGCCAGT	3480
OY	4643	CCAAACATGATTTATTAATCAAGGTCAATCTGAACATTTGAATATCAAAAATTCGACAG	4702
Db	3481	CCAAACATGATTTATTAATCAAGGTCAATCTGAACATTTGAATATCAAAAATTCGACAG	3540
OY	4703	TGAAAGATGCGAAGGACTCGGTGAAGCGGGTGAACGACATGAACGCGTGTTCGCCCGG	4762
Db	3541	TGAAAGATGCGAAGGACTCGGTGAAGCGGGTGAACGACATGAACGCGTGTTCGCCCGG	3600
OY	4763	CTTCGATATGTATCCCTCAGACATTCGCAACAGCTCCCGCATTCGAGCAAGCTGAG	4822
Db	3601	CTTCGATATGTATCCCTCAGACATTCGCAACAGCTCCCGCATTCGAGCAAGCTGAG	3660
OY	4823	ACGCTGTGTTCCTACATCCAGATGAACCGCAATTAATCTGTGCTATCAACAGACATCG	4882
Db	3661	ACGCTGTGTTCCTACATCCAGATGAACCGCAATTAATCTGTGCTATCAACAGACATCG	3720
OY	4883	GAGGTAAACAACATCTCAATGCGCGCAATACCTTCGCGGCTCAGAGAGAGAGACATTCG	4942
Db	3721	GAGGTAAACAACATCTCAATGCGCGCAATACCTTCGCGGCTCAGAGAGAGAGACATTCG	3780
OY	4943	TCAATGAGCGCTCGCTGCGCGGCAAGATGGCCGCGCAACCTCGCTGCGCGAGCGGATCGA	5002
Db	3781	TCAATGAGCGCTCGCTGCGCGGCAAGATGGCCGCGCAACCTCGCTGCGCGAGCGGATCGA	3840
OY	5003	GCGTGGAGCATCTCCAGAAAGAGAGATGTGCGGTGGCGGCTTCGTCTGGGTCCAGCTCT	5062
Db	3841	GCGTGGAGCATCTCCAGAAAGAGAGATGTGCGGTGGCGGCTTCGTCTGGGTCCAGCTCT	3900
OY	5063	GA 5064	
Db	3901	GA 3902	
RESULT 4			
AADD4951			
ID	AADD4951	standard; DNA; 3902 BP.	
XX	AADD4951:		
XX	13-DEC-2002	(first entry)	
XX	REB	(rice endosperm bZIP protein) transcription factor DNA.	
XX	Plant-derived food; flour; malt; monooct seed; milk protein.		
XX	infant formula; ds.		
XX	Unidentified.		





|||||  
Db 3781 TCAATGAGCGCTGTGCTGCCGCAAGATTGGCCGGCCAGACCTCCGTCGACAGGGGTGGCGA 3840  
QY 5003 GCGTGGAGCACTTCACAGAGAGATGCGGTGGCCCGGCTTCGCTGGCTCAGCTCCT 5062  
Db 3841 GCGTGGAGCACTTCACAGAGAGATGCGGTGGCCCGGCTTCGCTGGCTCAGCTCCT 3900  
QY 5063 GA 5064  
Db 3901 GA 3902  
  
RESULT 5  
ABS6544  
ID ABS6544 standard: DNA: 3902 BP.  
XX  
AC ABS6544;  
XX  
DT 29-NOV-2002 (first entry)  
XX  
DE Reb transcription factor sequence.  
XX  
KW Transformed plant; heterologous transcription factor; transgenic plant;  
KW seed protein; protein expression; gene expression control sequence;  
XX  
OS Synthetic.  
XX  
PN NC0200264750-A2.  
XX  
PD 22-APR-2002.  
XX  
PF 14-FEB-2002; 2002MO-US04909.  
XX  
PR 14-FEB-2001; 2001US-269189P.  
PR 14-FEB-2001; 2001US-269189P.  
PR 02-MAY-2001; 2001US-0847232.  
XX  
PA (VENT-) VENTRIA BIOSCIENCE.  
XX  
PI Huang N, Yang D;  
XX  
DR WPI; 2002-657592/70.  
XX  
PT Producing heterologous polypeptide in plant grain, by culturing  
PT transformed plant to form a grain-producing transforming plant, and  
PT recovering transgenic grains containing polypeptide from transformed  
PT plant -  
XX  
PS Exampe 15; Page 124-125; 230pp; English.  
XX  
CC The invention describes a method of producing a heterologous polypeptide  
CC (1) in a grain of a plant, comprising culturing a transformed plant (P1)  
CC comprising a first chimeric gene, and optionally, at least one  
CC heterologous transcription factor that is capable of enhancing the  
CC expression of the chimeric gene, to form a grain producing transforming  
CC plant (P2), and recovering transgenic grains containing (1) from P2.  
CC The method is useful for producing heterologous polypeptide in a grain  
CC of a plant. (1) is a non-plant storage, human or non-human animal, milk  
CC or other than a milk polypeptide, antibodies, cytokines, lymphokines,  
CC chemokines, hormones, growth factors, coagulation factors,  
CC anti-infectives, or cytotoxins, or anti-inflammatory molecule or  
CC intestinal trefoil factor (ITF) or its active fragment. Preferably, (1)  
CC is lactoferrin, lysozyme, lactoferricin, ITF, epidermal growth factor  
CC (EGF), keratinocyte growth factor (KGF), insulin-like growth factor I  
CC (IGF-1), lactohedin, kappa-casein, haptocorrin, lactoperoxidase,  
CC alpha-1-antitrypsin, immunoglobulins, alpha-lactalbumin,  
CC beta-lactoglobulin, alpha-casein, beta-casein, albumin, fibrinogen or  
CC protease inhibitor. This sequence represents a gene expression control  
CC sequence used in the method of producing a transgene plant.  
XX  
SQ Sequence 3902 BP; 975 A; 824 C; 937 G; 1166 T; 0 other;

Query Match 62.7%; Score 3902; DB 24; Length 3902;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3902; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1163 ATGGAGCGGATGTTCTCCGTGAGAGATCTCCGACCATTTCTGGTCCGCTCCGCCG 1222  
Db 1 ATGGAGCGGATGTTCTCCGTGAGAGATCTCCGACCATTTCTGGTCCGCTCCGCCG 60  
QY 1223 CCGGATGCG 1282  
Db 61 CCGGATGCG 120  
QY 1283 GGTGTGTAGCGGAGGCG 1342  
Db 121 GGTGTGTAGCGGAGGCG 180  
QY 1343 TACTTCCAGAGTTTCTGAGAGAGCGGTCTGATAGCCCGTCCGAACTTACCCG 1402  
Db 181 TACTTCCAGAGTTTCTGAGAGAGCGGTCTGATAGCCCGTCCGAACTTACCCG 240  
QY 1403 AGGCGCGAAGCGGAGGAGATCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1462  
Db 241 AGGCGCGAAGCGGAGGAGATCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300  
QY 1463 CCGGAGCTTCG 1522  
Db 301 CCGGAGCTTCG 360  
QY 1523 ATGCTGAGAGAGAGTGGAGAGAGACCTCGCGCGCGCGCGCGCGCGCGCGCGCG 1582  
Db 361 ATGCTGAGAGAGTGGAGAGAGACCTCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420  
QY 1583 TTCTCCCCCTCTAGTACTGAGAGAGTCTACTAGATCGGCAATGCTAGTCTGTTTC 1642  
Db 421 TTCTCCCCCTCTAGTACTGAGAGAGTCTACTAGATCGGCAATGCTAGTCTGTTTC 480  
QY 1643 ATCGAATGTTTATAGTATTTAGATCGGCGATTTCTATAGACCAATGGCTCCATG 1702  
Db 481 ATCGAATGTTTATAGTATTTAGATCGGCGATTTCTATAGACCAATGGCTCCATG 540  
QY 1703 TGCATGCGCTGTGATGATGCGGTGCGGTGCTTCACACATAGATATAGGTTGTGG 1762  
Db 541 TGCATGCGCTGTGATGATGCGGTGCGGTGCTTCACACATAGATATAGGTTGTGG 600  
QY 1763 TATGTACAAAGGAGATTCATAGACCTCGGATTTAGATTCGATATCGATCGACCTG 1822  
Db 601 TATGTACAAAGGAGATTCATAGACCTCGGATTTAGATTCGATATCGATCGACCTG 660  
QY 1823 CGAATGCGATGTATATGATGCGATGCTTACTAATCTCGAGACAGTGTGTTTGA 1882  
Db 661 CGAATGCGATGTATATGATGCGATGCTTACTAATCTCGAGACAGTGTGTTTGA 720  
QY 1883 ATTCGAATTTTGTATGAGGCGCTTAACATGCTATGCTTTTATGATGGCAATAT 1942  
Db 721 ATTCGAATTTTGTATGAGGCGCTTAACATGCTATGCTTTTATGATGGCAATAT 780  
QY 1943 GTGACTTTATGCAAGATTTGATGCGGATTCAGAGAACAAAGTGGGGGATTCGA 2002  
Db 781 GTGACTTTATGCAAGATTTGATGCGGATTCAGAGAACAAAGTGGGGGATTCGA 840  
QY 2003 CATACGAGTACCTGGCTAATACACATCATCTCATTTAATCATGCTAAATGCTAT 2062  
Db 841 CATACGAGTACCTGGCTAATACACATCATCTCATTTAATCATGCTAAATGCTAT 900  
QY 2063 TAAAGACCTTTAGCACTCTTATCTATTCACACATGAGTGAATAAATGAAGGGGAG 2122  
Db 901 TAAAGACCTTTAGCACTCTTATCTATTCACACATGAGTGAATAAATGAAGGGGAG 960  
QY 2123 CAGGGGGTATCCATGGTCCGATGGTGCAGGGGGGAGTCCCGCCGACCCAGC 2182  
Db 961 CAGGGGGTATCCATGGTCCGATGGTGCAGGGGGGAGTCCCGCCGACCCAGC 1020  
QY 2183 TTGAATCGCCCTGGAGTATGCTATAGCTGTCAAGCCATTTCTAGTGTCTGTGTAG 2242









XX WO200231154-A1.  
 PN 18-APR-2002.  
 XX 11-OCT-2001; 2001WO-JP08936.  
 XX 11-OCT-2000; 2000JP-0311295.  
 XX (NAG-) NAT INST AGROBIOLOGICAL SCI.  
 PA (BIO-) BIO-ORIENTED TECHNOLOGY RES. ADVANCEMENT.  
 XX Takaiwa F, Onodera Y;  
 PI WPI: 2002-372276/40.  
 DR P-PSDB; AAO14843.  
 XX Rice seed-originated bZIP type transcription factors regulating  
 PT expression of rice storage protein with binding activity to GCN4 motif,  
 PT useful in constructing new breeds of plants to produce seeds with  
 PT higher nutrition -  
 XX  
 PS Claim 1; Page 68-72; 124pp; Japanese.  
 CC The invention comprises the amino acid and coding sequences of rice seed  
 CC b-ZIP type transcription factors (RISB21, RISB24 and RISB25). The DNA and  
 CC protein sequences of the rice seed b-ZIP transcription factors are useful  
 CC in constructing new breeds of plants (e.g. rice) - to produce seeds with  
 CC higher nutrition and denser protein storage. The present cDNA sequence  
 CC encodes the rice seed b-zipper 1 (RISB21) protein.  
 XX  
 XX Sequence 1751 BP; 436 A; 425 C; 509 G; 381 T; 0 other;  
 SQ  
 Query Match 2.2%; Score 138.8; DB 24; Length 1751;  
 Best Local Similarity 62.4%; Pred. No. 5.1e-23;  
 Matches 272; Conservative 0; Mismatches 152; Indels 12; Gaps 3;  
 QY 1151 TCCGATTCGGAGTGGAGGCGGGTTCCTCGTGAAGAGATCCGACCCATCTGGGTC 1210  
 DB 199 TCCGAGAGCGCATGAGGACACGTTGCGCTGACGAGATCCCGACCCGCTGGGCT 258  
 QY 1211 CCGCCTCCGCGCCGCGCAGTGGCGGCGCGCCCGCAGCAGCAGGCGCGCGCGCTG-GC 1269  
 DB 259 CCGCGCGCGCGGTGGCAGCGCGGCGCGCCCGCAGATGATGACTGGCGCGGTAGC 318  
 QY 1270 TTCGGAGAGTGTGTGTCTAGCGGCGGCGCGCGCGCGGAGACGCGATGACCGGTG 1329  
 DB 319 GCGGCGGCGGTGTGTGAGCGGTCCCGTGGGAGTGAACCTCGAGAGGTTCTGGAGAG 378  
 QY 1330 CCGGTGAGAGTGTACTTCC-----AGAAGTTCTGGAGAGGCGGCTGCTGATAGCCCC 1384  
 DB 379 CTCGACGGCGTCCCTGCGACCGCGGCGGAGCCCGAGCGCGCGCATTTAACCTAGCCCG 438  
 QY 1385 GTCCGGAACCTAGCCCGAGGCGCGAGCGGAGGATCAGGCGCGCGAGAGGGGTGTG 1444  
 DB 439 ATGCCGCGCGCGCGGAGGCGGCGGCTGAGTGGGCTACCGCGCATCTGTGAG 498  
 QY 1445 CCGGTGATGTTAAGCAGCGCGCAGTCTCG-----CGGCGCGCGCAGCAGCGCGGTG 1498  
 DB 499 GCGGTGGGGGTGATGCCCATGCCCGGCGCGCTTCCGCGCGCGCGCGAGCGCGCG 558  
 QY 1499 GTGAGACCCGTGAGAGTCAACGCGATGCTGAGAGCAGAGTGGAGAGGACCTGCCGCG 1558  
 DB 559 ATGAGACCCGTGAGAGTCAACGCGATGCTGAGAGCGAGAGCTGACGAGAGCATCTGCCAC 618  
 QY 1559 GTGCCATGTGAGAGG 1574  
 DB 619 GTGCCATGTGAGAGG 634

XX AAV12743;  
 AC 05-MAY-1998 (first entry)  
 DT RBEL transcription controlling factor #39 (RGEF2) coding sequence.  
 XX RBEL; rice starch branching enzyme; transcription controlling factor; ss.  
 XX  
 XX Oryza sativa.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 17..913  
 FT /\*tag= a  
 XX  
 PN JP10014575-A.  
 XX 20-JAN-1998.  
 PD 28-JUN-1996; 96JP-0169246.  
 XX 28-JUN-1996; 96JP-0169246.  
 XX  
 PR 28-JUN-1996; 96JP-0169246.  
 XX  
 PA (MITS-) MITSUI GYOSAI SHOKUBUTSU BIO KENKYUSHO.  
 PA (MITR ) MITSUI TOHITSU CHEM INC.  
 DR WPI; 1998-138237/13.  
 DR P-PSDB; AAW1153.  
 XX  
 PT Transcription controlling factor in rice - comprises starch  
 PT branching enzyme gene, allows rice to be improved in quality and  
 PT quantity  
 XX  
 PS Claim 3; Page 7-8; 11pp; Japanese.  
 CC This sequence represents the DNA encoding a transcription controlling  
 CC factor of the invention, designated #39 (RGEF2). The transcription  
 CC controlling factor in rice promotes transcription of the RBEL gene, use  
 CC of the transcription factor enables rice to be improved in quality and  
 CC quantity.  
 XX  
 SQ Sequence 1086 BP; 321 A; 223 C; 283 G; 259 T; 0 other;  
 Query Match 1.5%; Score 92.6; DB 19; Length 1086;  
 Best Local Similarity 81.7%; Pred. No. 6.4e-12;  
 Matches 107; Conservative 0; Mismatches 24; Indels 0; Gaps 0;  
 QY 4370 CAGGTATCGCAATTAAGATCGAAGACTCTCGCTGTAAAGCGTCTTCTGATGTTAAC 4429  
 DB 401 CAGGTATCACTATTAAAGGTTGAAACTCTCTCTGTGAGCGCTTGTGATGCAAAAT 460  
 QY 4430 CAGAGTACATAGATGCTGCTGTGACATAGAGTCTAAAGCAGATGTTAGACCTTG 4489  
 DB 461 CAGAGTACATAGTCT 520  
 QY 4490 AGAGCAAAAGT 4500  
 DB 521 AGAGCAAAAGT 531  
 RESULT 9  
 AAD45324  
 ID AAD45324 standard; DNA; 1314 BP.  
 XX  
 AC AAD45324;  
 XX  
 DT 27-DEC-2002 (first entry)  
 XX  
 XX Maize opaque 2 (O2) transcription factor DNA.  
 XX  
 KW Maize; feed improvement; plant-derived feed; antibiotic; additive;  
 KW anti-microbial; poultry; alpha-1-antitrypsin; flour; malt; opaque 2;  
 KW O2 transcription factor; ds.

XX Zea mays.  
 OS  
 XX  
 PN WO200263975-A2.  
 XX  
 PD 22-AUG-2002.  
 XX  
 PF 14-FEB-2002; 2002WO-US04919.  
 XX  
 PR 14-FEB-2001; 2001US-269188P.  
 PR 02-MAY-2001; 2001US-0847232.  
 PA  
 XX (VENT-) VENTRIA BIOSCIENCE.  
 PI  
 XX Huang N, Rodriguez RL, Hagie FE;  
 DR WPI; 2002-682708/73.  
 XX  
 XX Improved feed for production animals, comprising plant-derived feed  
 PT ingredients, and seed composition containing flour, extract, or malt  
 PT from mature monooct seeds and heterologous seed-produced anti-microbial  
 PT proteins  
 XX  
 XX Disclosure; Page 158; 175pp; English.  
 XX  
 CC The invention relates to improved feed for production animals,  
 CC comprising one or more plant-derived feed ingredients, substantially  
 CC unsupplemented with small-molecule antibiotics and as an additive a  
 CC seed composition containing a flour, extract or malt obtained from  
 CC mature monooct seeds and one or more heterologous seed-produced anti-  
 CC microbial proteins in substantially unpurified form. The invention  
 CC is useful as a feed for production animals such as poultry and hoofed  
 CC farm animals. The present sequence is maize opaque 2 (O2) transcription  
 CC factor DNA. This sequence is used in the invention.  
 XX  
 XX Sequence 1314 BP; 333 A; 358 C; 396 G; 227 T; 0 other;  
 SQ  
 Query Match 1.4%; Score 85.4; DB 24; Length 1314;  
 Best Local Similarity 55.3%; Pred. No. 3.9e-10;  
 Matches 236; Conservative 0; Mismatches 176; Indels 15; Gaps 3;  
 QY 1163 ATGAGCGGGGTGTCCTCCGTGAGAGATCCGACCCATTGTGGTCCGCCCTCCGCCG 1222  
 DB 1 ATGAGCACGCTCATCTCAATGAGAGAGATCTCGGCGCTTCTGGAGCTGTACACCG 60  
 QY 1223 CCGCAGTCGCGCGCGGCCACGACGAGG-----CGCGCGCGCGCTTGGGGA 1276  
 DB 61 CCAGCGCCAGAGCCAGAGAGAGAGAGCGCTCCGTTAACCGGCTGCGGCAATGTC 120  
 QY 1277 GGTGGTGGTGTGTAAGGAGGCGGCGCGGCGGAGGAGCGCATGAACCGGTCGCG 1336  
 DB 121 ATGACGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180  
 QY 1337 GAGTGTACTTCCAGAAATTTCTGAGAGAGCGGTGCT--CGATAGCCCGCTCCGAAAC 1394  
 DB 181 GAGTGTACTTCTGAGAGGTACTAGAGAGAGAGAGCTCTGACGACAGCAGCAGCGCGCG 240  
 QY 1395 CTAAGCCGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1447  
 DB 241 GTGGTGTGGTGTGCGAATCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300  
 QY 1448 GTGATGTTAAGCAGCCGAGCTCTCGGCGCGGCGGAGAGCAGCAGCGGTGGTGAACCC 1507  
 DB 301 ATGGAAGAGGCGGTAACTATGCGCGCTGCGGTGGAGTAGTGTCCCTATAGTGAACCC 360  
 QY 1508 GTGAGTACAAAGCGATGCTGAAGCAGAAAGCTGGAAGAGAGAGAGAGAGAGAGAGAG 1567  
 DB 361 ATGAGTACAAAGCGATGCTGAAGCAGAAAGCTGGAAGAGAGAGAGAGAGAGAGAGAG 420  
 QY 1568 TGAAGG 1574  
 DB 421 TGAAGG 427

RESULT 10  
 AAD44949  
 ID AAD44949 standard; DNA; 1314 BP.  
 XX  
 AC AAD44949;  
 XX  
 DT 13-DEC-2002 (first entry)  
 XX  
 DE 02 (Opaque 2) transcription factor DNA.  
 XX  
 KW Plant-derived food; flour; malt; monooct seed; milk protein;  
 KW infant formula; ds.  
 XX  
 OS Unidentified.  
 XX  
 PN WO200264814-A2.  
 XX  
 PD 22-AUG-2002.  
 XX  
 PF 14-FEB-2002; 2002WO-US04921.  
 XX  
 PR 14-FEB-2001; 2001US-269199P.  
 PR 02-MAY-2001; 2001US-0847232.  
 XX  
 PA (VENT-) VENTRIA BIOSCIENCE.  
 PI  
 XX Huang N, Rodriguez RL, Hagie FE;  
 DR WPI; 2002-667011/71.  
 XX  
 XX New nutritionally enhanced food compositions, useful for improving  
 PT infant formula, or supplementing or enhancing the diet of infants,  
 PT particularly very-low birth weight infants  
 XX  
 XX Example 9; Page 111; 179pp; English.  
 PS  
 CC The invention relates to a nutritionally enhanced food comprising one  
 CC or more plant-derived food ingredients, and as an additive, a seed  
 CC composition containing a flour, extract, or malt obtained from mature  
 CC monooct seeds and one or more seed-produced human milk proteins in  
 CC substantially unpurified form. The nutritionally enhanced food and food  
 CC additive compositions are useful for improving infant formula. Infant  
 CC formulas containing the recombinant human milk proteins are useful in  
 CC supplementing or enhancing the diet of infants, particularly very-low  
 CC birth weight infants. The present sequence is a O2 (Opaque 2)  
 CC transcription factor DNA used in the invention.  
 XX  
 XX Sequence 1314 BP; 333 A; 358 C; 396 G; 227 T; 0 other;  
 SQ  
 Query Match 1.4%; Score 85.4; DB 24; Length 1314;  
 Best Local Similarity 55.3%; Pred. No. 3.9e-10;  
 Matches 236; Conservative 0; Mismatches 176; Indels 15; Gaps 3;  
 QY 1163 ATGAGCGGGGTGTCCTCCGTGAGAGATCCGACCCATTGTGGTCCGCCCTCCGCCG 1222  
 DB 1 ATGAGCACGCTCATCTCAATGAGAGAGATCTCGGCGCTTCTGGAGCTGTACACCG 60  
 QY 1223 CCGCAGTCGCGCGCGGCCACGACGAGG-----CGCGCGCGCGCTTGGGGA 1276  
 DB 61 CCAGCGCCAGAGCCAGAGAGAGAGAGAGCGCTCCGTTAACCGGATGCTGCGGCAATGTC 120  
 QY 1277 GGTGGTGGTGTGTAAGGAGGCGGCGGCGGAGGAGCGCATGAACCGGTCGCG 1336  
 DB 121 ATGACGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180  
 QY 1337 GAGTGTACTTCCAGAAATTTCTGAGAGAGCGGTGCT--CGATAGCCCGCTCCGAAAC 1394  
 DB 181 GAGTGTACTTCTGAGAGGTACTAGAGAGAGAGAGAGCTCTGACGACAGCAGCAGCGCGCG 240  
 QY 1395 CTAAGCCGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1447  
 DB 241 GTGGTGTGGTGTGCGAATCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300





Db 6381 TATACACACTTATATACATATATACCTATATATATAATAACACATATATATATATACC 6322

CC The invention relates to a composition comprising a human GDP-mannose  
CC 4,6-dehydratase (GM4, 6b) peptide. The peptide is useful for identifying  
CC GM4, 6b inhibitors. GM4, 6b inhibitors are useful for reducing inflammation  
CC in a mammalian subject and for treating or ameliorating diseases affected  
CC by the level of cellular fucosylation or diseases affected by the  
CC translocation of glycoconjugates. These diseases include arthritis,  
CC transplant rejection, asthma, sepsis, reperfusion injury, stroke or  
CC infection. The GM4, 6b peptide or a polynucleotide encoding it is also  
CC useful for manufacturing complex carbohydrates and as targets for  
CC screening small molecule antagonists of the activity of the enzyme. The  
CC polynucleotide is useful in developing an assay for defects in the  
CC enzyme, as well as in gene replacement therapy. Sequences  
CC ABX11942-ABX17944 and ABX11941-ABX33716 represent DNA molecules encoding  
CC human GM4, 6b peptides of the invention.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format directly from USPTO  
CC at [seqdata.uspto.gov/sequence.htm](http://seqdata.uspto.gov/sequence.htm).

XX Sequence 388 BP; 118 A; 82 C; 80 G; 108 T; 0 other;  
SQ Query Match 1.48; Score 84.2; DB 25; Length 388;  
Best Local Similarity 78.3%; Pred. No. 4.1e-10;  
Matches 101; Conservative 0; Mismatches 28; Indels 0; Gaps 0;  
QY 4372 GGTATCGCAATTAAAGTCGAGAACTCCCTGCTTTAAAGCGCTTGTGATGTTAACCA 4431  
DB 1 GGTTCCTCAATTAAAGGTGAAATTCACCTTCTTAAAGCGCTTACTGATGTTAGTCA 60  
QY 4432 GAATACATGATGCTGCTGTTGACATATGAGTGTCTTAAAGCAATGTTGAGACCTTGAG 4491  
DB 61 GAATACAGCGACTGCTGTTGACACAGAGTTTAAAGCTGATGTTGAACATTAAAG 120  
QY 4492 AGCAAGGT 4500  
DB 121 AACNAAGGT 129

Search completed: September 13, 2003, 19:02:19  
Job time : 1046.02 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Comphen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 13, 2003, 02:54:05 ; Search time 8198.38 Seconds  
(without alignments)  
18460.230 Million cell updates/sec

Title: US-09-847-232a-35  
Perfect score: 6227  
Sequence: 1 gttaccatcatacatacatta.....atracacgctgcaagctt 6227

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues  
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estum:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_huv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vtc:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_man:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rtd:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vtl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
C 1	735.6	11.8	827	14	CB654912 OSJNEC07M
C 2	667	10.7	801	14	CB669209 OSJNEC011
C 3	632	10.1	701	14	AU092336 AU092336
C 4	586.4	9.4	869	14	CB659203 OSJNEC15P

Result No.	Score	Query Match	Length DB	ID	Description
5	554.8	8.9	836	14	CB668872 OSJNEC160
6	545.8	8.8	760	14	CB662942 OSJNEC07G
7	533.8	8.6	822	14	CB659202 OSJNEC15P
8	531.4	8.5	680	14	CB641559 OSJNEC01E
9	530	8.5	845	14	CB654911 OSJNEC07M
10	516.8	8.3	832	14	CB669208 OSJNEC011
11	459.4	7.4	663	12	B1808598 D004E01 O
12	441.2	7.1	464	9	AU179335 AU179335
13	422.8	6.8	566	12	B1809493 F010E11 O
14	407.4	6.5	421	13	BQ908360 T006G11 O
15	354	5.7	813	14	CA757602 OE04A06-T
16	343.2	5.5	419	14	C26764 C26764
17	266.8	4.3	831	10	BG320189 Zm03_05e0
18	266.8	4.3	1545	11	AY109969 Zea mays
19	265.2	4.3	568	12	BM268814 MEST400-B
20	265.2	4.3	595	12	BM072951 MEST56-D0
21	256.6	4.1	404	28	BH783568 f2mb013f0
22	254.8	4.1	631	14	CD207949 Hs1_36_C1
23	253	4.1	670	9	AM563718 L61_248_E
24	253	4.1	684	10	BE597974 P11_67_G1
25	253	4.1	710	14	CD211880 Hs1_67_A1
26	243.2	3.9	1663	11	AY109970 Zea mays
27	235.4	3.8	586	12	BM080681 MEST113-A
28	229.2	3.7	508	12	BM074279 MEST83-G0
29	227.8	3.7	624	9	AM330755 707044A01
30	214	3.4	250	14	C26541 C26541
31	212.8	3.4	671	13	BQ160449 WHE817-1
32	209.4	3.4	707	12	BJ308621 BJ308621
33	208	3.3	722	13	B0801405 WHE8214.A
34	191.6	3.1	411	9	AT586550 486052H05
35	187	3.0	484	10	BE637338 WHE8187-1
36	184.6	3.0	595	13	BQ665560 HX04C15u
37	166.8	2.7	602	29	B2542906 OGAKX29TM
38	161.4	2.6	811	29	B2542900 OGAKX29TC
39	159	2.6	485	14	CA697089 W1K4_PK00
40	149.8	2.4	559	14	CA830444 1117006H1
41	149.8	2.4	584	13	BQ779065 946116P08
42	148.8	2.4	501	14	CA830463 1117007C0
43	148.6	2.4	521	13	B0097966 946122F06
44	143.4	2.3	890	29	B2753115 P0FEB09PD
45	142.4	2.3	540	14	CD211973 Hs1_67_A1

## ALIGNMENTS

RESULT 1  
CB654912/c 827 bp mRNA linear EST 09-Apr-2003  
LOCUS OSJNEC07M21.1 OSJNEC Oryza sativa (japonica cultivar-group) cDNA  
DEFINITION clone OSJNEC07M21.1 GI:29658637

ACCESSION CB654912  
VERSION CB654912.1 GI:29658637  
KEYWORDS EST.

SOURCE Oryza sativa (japonica cultivar-group)  
ORGANISM Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE 1 (bases 1 to 827)  
Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,  
Kudrna,D., Dean,R., Soderlund,C., Wang,R. and Wang,G.  
Large-scale identification of ESTs involved in the interaction  
between rice and Magnaporthe grisea

## JOURNAL COMMENT

Unpublished  
Contact: Rod Wing  
Arizona Genomics Institute  
University of Arizona  
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ  
85721-0088, USA  
Tel: 520 626 3967  
Fax: 520 621 9288  
Email: <http://genome.arizona.edu>



PCR Primers  
 FORWARD: gta aac cga cgg cca gtc  
 BACKWARD: gga aac agc tat gac cat g  
 Plate: 07 row: M column: 21  
 Seq primer: gga aac agc tat gac cat g.

## FEATURES

Location/Qualifiers

1. 827

/organism="Oryza sativa (japonica cultivar-group)"

/mol\_type="mRNA"

/cultivar="Nipponbare"

/db\_xref="taxon:39947"

/clone="OSJNE01111"

/tissue\_type="leaf"

/dev\_stage="3 week"

/lab\_host="DH10B"

/clone\_idb="OSJNEc"

/note="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2: XhoI; 6 hrs after inoculation with Rice Blast (C9240-1)"

## BASE COUNT

214 a 208 c 199 g 206 t

## ORIGIN

11.8%; Score 735.6; DB 14; Length 827;

Query Match Best Local Similarity 99.5%; Pred. No. 8e-152; Indels 0; Gaps 0;

Matches 738; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4700 AGGTGAAGATGGCAGAGAGACTGGTGAAGCGGGTGACAGCATGAACGGTTGTTCCG 4759  
 DB 750 AGGTGAAGATGGCAGAGAGACTGGTGAAGCGGGTGACAGCATGAACGGTTGTTCCG 691  
 QY 4760 CCGCTTCGATGATGATCATCCCTCAGCATGCGCATTCAGACAGCTCCCACTAAGCAAG 4819  
 DB 690 CCGCTTCGATGATGATCATCCCTCAGCATGCGCATTCAGACAGCTCCCACTAAGCAAG 631  
 QY 4820 CAGAGCTGCTGTTCCCATCAAGATGACCGGACCAATTCCTGCTACTAACAAGACA 4879  
 DB 630 CAGAGCTGCTGTTCCCATCAAGATGACCGGACCAATTCCTGCTACTAACAAGACA 571  
 QY 4880 TCGAGGTACACACACTACATGCCGACATACCTTTCTGGCTCAGAGAGAGAGACT 4939  
 DB 570 TCGAGGTACACACACTACATGCCGACATACCTTTCTGGCTCAGAGAGAGAGACT 511  
 QY 4940 TCGTCAATGGCGCTCGGCTGCGGCGGAGATTTGGCCGGCAGCTGCGTGCAGCGGGTGG 4999  
 DB 510 TCGTCAATGGCGCTCGGCTGCGGCGGAGATTTGGCCGGCAGCTGCGTGCAGCGGGTGG 451  
 QY 5000 CGAGCTGGAGCATCTCCAGAAAGAGATGTCCGGTGGCGGCTGCTGGGTGAGAGT 5059  
 DB 450 CGAGCTGGAGCATCTCCAGAAAGAGATGTCCGGTGGCGGCTGCTGGGTGAGAGT 391  
 QY 5060 CCGTGAAGCGGAACCCGAGAGCTGCTGGGTCTGAAGACACTGCGAGAGAGAAATGATG 5119  
 DB 390 CCGTGAAGCGGAACCCGAGAGCTGCTGGGTCTGAAGACACTGCGAGAGAGAAATGATG 331  
 QY 5120 ATTGAGACAGGCTAGACATGCTAATGCTGAGAGTTGATGATGTTGGTCTGCTGCTC 5179  
 DB 330 ATTGAGACAGGCTAGACATGCTAATGCTGAGAGTTGATGATGTTGGTCTGCTGCTC 271  
 QY 5180 GTCAATGTGCAATCTTTGTAAGGAGACCTCTTATACCTCTTCTTCTTAAGAGACTAG 5239  
 DB 270 GTCAATGTGCAATCTTTGTAAGGAGACCTCTTATACCTCTTCTTCTTAAGAGACTAG 211  
 QY 5240 TACCCCTTGGATCTGATGCTCTAATTAATACATTAACATTAAGCAAGTTCATGGTG 5299  
 DB 210 TACCCCTTGGATCTGATGCTCTAATTAATACATTAACATTAAGCAAGTTCATGGTG 151  
 QY 5300 TGAATGCGCTGCTCCCTAATTTGAAGACAGTATTCAGAGAGCTATTCATTCATTCAA 5359  
 DB 150 TGAATGCGCTGCTCCCTAATTTGAAGACAGTATTCAGAGAGCTATTCATTCATTCAA 91  
 QY 5360 TAAATATTTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5419  
 DB 90 TAAATATTTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 31

QY 5420 TCACTACTTTTGACCTGTTCA 5441  
 DB 30 TCACTACTTTTGACCTGTTCA 9

## RESULT 2

CB669209/c

LOCUS

DEFINITION

OSJNE01111.1 OSJNE Oryza sativa (japonica cultivar-group) cDNA

ACCESSION

CB669209

KEYWORDS

SOURCE

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrhartoideae; Oryzaceae; Oryza.

1 (bases 1 to 801)

Jantauriyarat,C., Lu,G., Gowda,M., Hatfield,V., Zhou,B., Mazur,E.,

Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.

Large-scale identification of ESTs involved in the interaction

between rice and Magnaporthe grisea

Unpublished

CONTACT: Rod Wing

Arizona Genomics Institute

University of Arizona

Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ

85721-0088, USA

Tel: 520 626 3967

Fax: 520 621 9288

Email: http://genome.arizona.edu

PCR Primers

FORWARD: gta aac cga cgg cca gtc

BACKWARD: gga aac agc tat gac cat g

Plate: 01 row: I column: 11

Seq primer: gga aac agc tat gac cat g.

Location/Qualifiers

1. 801

/organism="Oryza sativa (japonica cultivar-group)"

/mol\_type="mRNA"

/cultivar="Nipponbare"

/db\_xref="taxon:39947"

/clone="OSJNE01111"

/tissue\_type="leaf"

/dev\_stage="3 week"

/lab\_host="DH10B"

/clone\_idb="OSJNEc"

/note="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2: XhoI; 24 hrs after inoculation with Rice Blast (70-15)"

## BASE COUNT

191 a 211 c 192 g 207 t

## ORIGIN

Query Match

Best Local Similarity 10.7%; Score 667; DB 14; Length 801;

Matches 670; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4700 AGGTGAAGATGGCAGAGAGACTGGTGAAGCGGGTGACAGCATGAACGGTTGTTCCG 4759  
 DB 721 AGGTGAAGATGGCAGAGAGACTGGTGAAGCGGGTGACAGCATGAACGGTTGTTCCG 662  
 QY 4760 CCGCTTCGATGATGATCATCCCTCAGCATGCGCATTCAGACAGCTCCCACTAAGCAAGT 4819  
 DB 661 CCGCTTCGATGATGATCATCCCTCAGCATGCGCATTCAGACAGCTCCCACTAAGCAAGT 602  
 QY 4820 CAGAGCTGCTGTTCCCATCAAGATGACCGGACCAATTCCTGCTACTAACAAGACA 4879  
 DB 601 CAGAGCTGCTGTTCCCATCAAGATGACCGGACCAATTCCTGCTACTAACAAGACA 542  
 QY 4880 TCGAGGTACACACACTACATGCCGACATACCTTTCTGGCTCAGAGAGAGAGAGACT 4939  
 DB 541 TCGAGGTACACACACTACATGCCGACATACCTTTCTGGCTCAGAGAGAGAGAGACT 482

QY 4940 TCGTCAATGCGGCTCTGGCTGCGCGGACAGATTGGCCGAGCCTGCTGCAAGCGGGTGG 4999  
 |||||||  
 Db 481 TCGTCAATGCGGCTCTGGCTGCGCGGACAGATTGGCGGACAGCCTGCTGCAAGCGGGTGG 422  
 |||||||  
 QY 5000 CGAGCTGGAGCATCTCCAGAGAGATGTCGGTGGGCGGCTTGTGCTGGGTGACGT 5059  
 |||||||  
 Db 421 CGAGCTGGAGCATCTCCAGAGAGATGTCGGTGGGCGGCTTGTGCTGGGTGACGT 362  
 |||||||  
 QY 5060 CCTGAGACCGAAACCCAGAGCTCTCGGTTGCTTGAAGACACTGCGAGCAGGAATGATG 5119  
 |||||||  
 Db 361 CCTGAGACCGAAACCCAGAGCTCTCGGTTGCTTGAAGACACTGCGAGCAGGAATGATG 302  
 |||||||  
 QY 5120 ATTGACAGAGCGGTAGACATTGCTTAATGCTGTGAGGTGATGATGTTGGTGGCTGTC 5179  
 |||||||  
 Db 301 ATTGACAGAGCGGTAGACATTGCTTAATGCTGTGAGGTGATGATGTTGGTGGCTGTC 242  
 |||||||  
 QY 5180 GTCATTGTGCAATCTTTGTAAGGAGACCTGTAAGTACCTGCTTCTTCTAAAGGACTAG 5239  
 |||||||  
 Db 241 GTCATTGTGCAATCTTTGTAAGGAGACCTGTAAGTACCTGCTTCTTCTAAAGGACTAG 182  
 |||||||  
 QY 5240 TACCCCTTGTGATCTGATGCTTAAATTAATTAATACACATTAGCCAAATGTCATTGCTG 5299  
 |||||||  
 Db 181 TACCCCTTGTGATCTGATGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 122  
 |||||||  
 QY 5300 TGATGGCTGCTCCCTTAATTGACAGCTGATTTGACGAGCTGCTATGCTATCA 5359  
 |||||||  
 Db 121 TGATGGCTGCTCCCTTAATTGACAGCTGATTTGACGAGCTGCTATGCTATCA 62  
 |||||||  
 QY 5360 TAAATATTGATCGA 5374  
 |||||||  
 Db 61 TAAATATTGATCGA 47  
 |||||||

RESULT 3  
 A0092336 701 bp mRNA linear EST 03-APR-2002  
 LOCUS A0092336  
 DEFINITION clone Cl2552, mRNA sequence.  
 ACCESSION A0092336  
 VERSION A0092336.1 GI:8527521  
 KEYWORDS EST.  
 SOURCE Oryza sativa (japonica cultivar-group)  
 ORGANISM Oryza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.  
 REFERENCE 1 (bases 1 to 701)  
 AUTHORS Sasaki,T. and Yamamoto,K.  
 TITLE Rice cDNA from callus (2000)  
 JOURNAL Unpublished  
 COMMENT Contact: Takuji Sasaki  
 National Institute of Agrobiological Resources  
 Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki  
 305-8602, Japan  
 Tel: 81-298-38-7441  
 Fax: 81-298-38-7468  
 Email: tsasaki@abrr.affrc.go.jp, URL: http://rnp.dna.affrc.go.jp/  
 PROJECT -"RNP"

FEATURES  
 source  
 1..701  
 Location/Qualifiers  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /mol\_type="mRNA"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:39947"  
 /clone="Cl2552"  
 /clone\_1db="Rice callus"  
 /note="Vector: Bluescript II SK+, Site\_1: SalI; Site\_2:  
 NotI; cDNA prepared from rice callus mRNA by using  
 oligo(dT) as a primer and ligating to the SalI-NotI site  
 of Bluescript II SK+ phagemid."  
 BASE COUNT 178 a 172 c 179 g 170 t 2 others  
 ORIGIN

Query Match 10.1%; Score 632; DB 9; Length 701;  
 Best Local Similarity 98.8%; Pred. No. 6,9e-129;  
 Matches 657; Conservative 0; Mismatches 6; Indels 2; Gaps 2;

QY 4700 AGGTGAAGATGCGAGAGGACTGCTGGAAGCGGGTGACAGGCAATGACCGTTGTTCCG 4759  
 |||||||  
 Db 27 AGGTGAAGATGCGAGAGGACTGCTGGAAGCGGGTGACAGGCAATGACCGTTGTTCCG 86  
 |||||||  
 QY 4760 CCGCTTGTGATGATGATATCCCTCAGCATGCCATCAACAGCTCCCATCTGACAGCACT 4819  
 |||||||  
 Db 87 CCGCTTGTGATGATGATATCCCTCAGCATGCCATCAACAGCTCCCATCTGACAGCACT 146  
 |||||||  
 QY 4820 CAGAGCTGCTGTTTCCATCCAGATGACCCGAAATTAATCTTGTGCTACTACAGACA 4879  
 |||||||  
 Db 147 CAGAGCTGCTGTTTCCATCCAGATGACCCGAAATTAATCTTGTGCTACTACAGACA 205  
 |||||||  
 QY 4880 TCGGAGGTAAACAACACTATGATCCCGACATACCTTGTGCTGCTGAGAGAGAGAGACT 4939  
 |||||||  
 Db 206 TCGGAGGTAAACAACACTATGATCCCGACATACCTTGTGCTGCTGAGAGAGAGAGACT 264  
 |||||||  
 QY 4940 TCGTCAATGCGGCTCTGGCTGCGCGGACAGATTGGCCGAGCCTGCTGCAAGCGGGTGG 4999  
 |||||||  
 Db 265 TCGTCAATGCGGCTCTGGCTGCGCGGACAGATTGGCCGAGCCTGCTGCAAGCGGGTGG 324  
 |||||||  
 QY 5000 CGAGCTGGAGCATCTCCAGAGAGATGTCGGTGGGCGGCTTGTGCTGGGTGACGT 5059  
 |||||||  
 Db 325 CGAGCTGGAGCATCTCCAGAGAGATGTCGGTGGGCGGCTTGTGCTGGGTGACGT 384  
 |||||||  
 QY 5060 CCTGAGACCGAAACCCAGAGCTCTCGGTTGCTTGAAGACACTGCGAGCAGGAATGATG 5119  
 |||||||  
 Db 385 CCTGAGACCGAAACCCAGAGCTCTCGGTTGCTTGAAGACACTGCGAGCAGGAATGATG 444  
 |||||||  
 QY 5120 ATTGACAGAGCGGTAGACATTGCTTAATGCTGTGAGGTGATGATGTTGGTGGCTGTC 5179  
 |||||||  
 Db 445 ATTGACAGAGCGGTAGACATTGCTTAATGCTGTGAGGTGATGATGTTGGTGGCTGTC 504  
 |||||||  
 QY 5180 GTCATTGTGCAATCTTTGTAAGGAGACCTGTAAGTACCTGCTTCTTCTAAAGGACTAG 5239  
 |||||||  
 Db 505 GTCATTGTGCAATCTTTGTAAGGAGACCTGTAAGTACCTGCTTCTTCTAAAGGACTAG 564  
 |||||||  
 QY 5240 TACCCCTTGTGATCTGATGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 5299  
 |||||||  
 Db 565 TACCCCTTGTGATCTGATGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 624  
 |||||||  
 QY 5300 TGATGGCTGCTCCCTTAATTGACAGCTGATTTGACGAGCTGCTATGCTATCA 5359  
 |||||||  
 Db 625 TGATGGCTGCTCCCTTAATTGACAGCTGATTTGACGAGCTGCTATGCTATCA 684  
 |||||||  
 QY 5360 TATA 5364  
 || |  
 Db 685 AAAA 689  
 || |

RESULT 4  
 CB659203 869 bp mRNA linear EST 09-APR-2003  
 LOCUS CB659203/c  
 DEFINITION clone OSJNBCL5P20 3', mRNA sequence.  
 ACCESSION CB659203  
 VERSION CB659203.1 GI:29662928  
 KEYWORDS EST.  
 SOURCE Oryza sativa (japonica cultivar-group)  
 ORGANISM Oryza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.  
 REFERENCE 1 (bases 1 to 869)  
 AUTHORS Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,F.,  
 Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.  
 TITLE Large-scale identification of ESTs involved in the interaction  
 between rice and Magnaporthe oryzae  
 JOURNAL Unpublished  
 COMMENT Contact: Rod Wing

Arizona Genomics Institute  
University of Arizona  
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ  
85721-0088, USA  
Tel: 520 626 3967  
Fax: 520 621 9288  
Email: <http://genome.arizona.edu>

FOR Primers  
FORWARD: gta aac cga cgg cca gty  
BACKWARD: gga aac agc tat gac cat g  
Plate: 15 row: P column: 20  
Seq primer: gga aac agc tat gac cat g.

## FEATURES

Source

1. 869  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="OSJNCL15P20"  
/tissue\_type="Leaf"  
/dev\_stage="3 week"  
/lab\_host="DH10B"  
/clone\_id="OSJNCL1"  
/note="Vector: Bluescript II KS +; Site 1: EcoRI; Site 2:  
XhoI; 5 hrs after inoculation with Rice Blast (C9240-1)"  
BASE COUNT 192 a 237 c 209 g 231 t  
ORIGIN

## Query Match

Best Local Similarity 99.4%; Score 586.4; DB 14; Length 869;  
Matches 587; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

4700 AGGTGAAGATGCGAGAGAGACTCGGTGAAGCGGAGCAAGCATGAAGCGGTTTCCCG 4759  
|||||  
592 AGGTGAAGATGCGAGAGAGACTCGGTGAAGCGGAGCAAGCATGAAGCGGTTTCCCG 533  
|||||  
4760 CGGCTTCGATATGTCATCCCTCAGCATCCATTCACAGAGTCCCATCTGAAGAAAGT 4819  
|||||  
532 CGGCTTCGATATGTCATCCCTCAGCATCCATTCACAGAGTCCCATCTGAAGAAAGT 473  
|||||  
4820 CAGAGCGTGTGTCCATCCAAAGATGACCCGCAACATTAATCTCGTACTAACAAGACA 4879  
|||||  
472 CAGAGCGTGTGTCCATCCAAAGATGACCCGCAACATTAATCTCGTACTAACAAGACA 413  
|||||  
4880 TCGAGGTAACACAACTACATGCGGACATACCTCTCGGCTCAGAGAGAGAGACT 4939  
|||||  
412 TCGAGGTAACACAACTACATGCGGACATACCTCTCGGCTCAGAGAGAGAGACT 353  
|||||  
4940 TCGTCAATGCGCTCTGCGTCCCGGCAAGATTGGCCGAGCCTCGCTGCAAGCGGCTGG 4999  
|||||  
352 TCGTCAATGCGCTCTGCGTCCCGGCAAGATTGGCCGAGCCTCGCTGCAAGCGGCTGG 293  
|||||  
5000 CAGAGCGTGTGTCCATCCAAAGATGACCCGCAACATTAATCTCGTACTAACAAGACA 5059  
|||||  
292 CAGAGCGTGTGTCCATCCAAAGATGACCCGCAACATTAATCTCGTACTAACAAGACA 233  
|||||  
5060 CCGAGACCGAAACCCAGAGCTCTCGGTTCTGAAGACACTGCGAGAGAGAAATGATG 5119  
|||||  
232 CCGAGACCGAAACCCAGAGCTCTCGGTTCTGAAGACACTGCGAGAGAGAAATGATG 173  
|||||  
5120 AATGGACAGGCGTAGACATTGCTAATGCTGTAGAGTGTGATGTTGTTGCTGCTGCTC 5179  
|||||  
172 AATGGACAGGCGTAGACATTGCTAATGCTGTAGAGTGTGATGTTGTTGCTGCTGCTC 113  
|||||  
5180 GTCATGTCGATCTCTTTGTAGAGGAGACCTCTAGTACCTCTCTCTTAAGAGCTAG 5239  
|||||  
112 GTCATGTCGATCTCTTTGTAGAGGAGACCTCTAGTACCTCTCTCTTAAGAGCTAG 53  
|||||  
5240 TACCCCTGTGATCTCATGCTCTTAATTAATCAATTAAGACATTAAGCCAAA 5287  
|||||  
52 TACCCCTGTGATCTCATGCTCTTAATTAATCAATTAAGACATTAAGCCAAA 5

## RESULT 5

CB668872

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

Source

CB668872 836 bp mRNA linear EST 09-APR-2003  
OSJNED16024.1 OSJNED Oryza sativa (japonica cultivar-group) cDNA  
clone OSJNED16024.3, mRNA sequence.  
CB668872 GI:29672597  
EST.  
Oryza sativa (japonica cultivar-group)  
Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Euphorbiaceae; Oryzaeae; Oryza.  
1 (bases 1 to 836)  
Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,  
Kudrna,D., Deap,R., Soderlund,C., Wing,R. and Wang,G.  
large-scale identification of ESTs involved in the interaction  
between rice and Magnaporthe grisea  
Unpublished  
Contact: Rod Wing  
Arizona Genomics Institute  
University of Arizona  
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ  
85721-0088, USA  
Tel: 520 626 3967  
Fax: 520 621 9288  
Email: <http://genome.arizona.edu>  
PCR Primers  
FORWARD: gta aac cga cgg cca gty  
BACKWARD: gga aac agc tat gac cat g  
Plate: 16 row: O column: 24  
Seq primer: gga aac agc tat gac cat g.  
Location/Qualifiers  
1. 836  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="OSJNED16024"  
/tissue\_type="Leaf"  
/dev\_stage="3 week"  
/lab\_host="DH10B"  
/clone\_id="OSJNED"  
/note="Vector: Bluescript II KS +; Site 1: EcoRI; Site 2:  
XhoI; 24 hrs after inoculation with Rice Blast (C9240-1)"  
BASE COUNT 262 a 200 c 175 g 199 t  
ORIGIN

Query Match . 8.9%; Score 554.8; DB 14; Length 836;  
Best Local Similarity 95.4%; Pred. No. 9.4e-112;  
Matches 582; Conservative 0; Mismatches 27; Indels 1; Gaps 1;

5555 TGTCAATTTCTAATCAAAATTAGTGAATGATTTTCACCCAGATCCATCAAGTTCAT 5614  
|||||  
42 TTTTCTTCAATATCAATTAATGATGATGATTTTCACCCATATCCATCCATGATCAT 101  
|||||  
5615 CATTAAAGCAAGGTGTATTAATGTAATTTTCTGTTAGAGAAAAAAGAGAGC 5674  
|||||  
102 CATTAAAGCAAGGTGTATTAATGTAATTTTCTGTTAGAGAAAAAAGAGAGC 161  
|||||  
5675 TTTTATTAAGATTCACCGGTGGGTGTGAACAAATTAATCAATGATGATGATCCCGT 5734  
|||||  
162 CTTTATTAATTAAGTTCACCGGTGGGTGTGAACAAATTAATCAATGATGATGATCCCGT 221  
|||||  
5735 AAGGCAAGCCTAGCTAGACAAATAAGCAATTAATCCGTTATCAACCAACAAAGCTT 5794  
|||||  
222 AAGGCAAGCCTAGCTAGACAAATAAGCAATTAATCCGTTATCAACCAACAAAGCTT 261  
|||||  
5795 GCGCAAGGCGCAATAGGCAAGGAGCTCATGCTTGGGCGCAAGAAACGAATCAAGTG 5854  
|||||  
282 GCGCAAGGCGCTCAATAGGCAAGGAGCTCATGCTTGGGCGCAAGAAACGAATCAAGTG 341  
|||||  
5855 ATACATTGGCAGGAAACCAAAAAGAGG-CCATCATATCAATTCATTCATCAAGCGGC 5913

```

|||||
Db 342 ATACATGTGGCAGGAGACCAACAAAAGAGCCCATCCAAATCCATCCATCCAAAGCGGC 401
OY 5914 ATGAGAGACAGAGATGATTCACAGCTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5973
Db 402 ATGAGAGACAGAGATGATTCACAGCTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 461
OY 5974 ACTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 6033
Db 462 ACTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 521
OY 6034 TGTGGAGATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 6093
Db 522 TGTGGAGATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 581
OY 6094 CGAATCATCGCAGCTAAAGAGACCAATATATATATATATATATATATATATATATAT 6153
Db 582 CGAATCATCGCAGCTAAAGAGACCAATATATATATATATATATATATATATATATAT 641
OY 6154 TGCATTAATAC 6163
Db 642 TGTCCAGATC 651

RESULT 6
CB662942 760 bp mRNA linear EST 09-APR-2003
LOCUS OSJNEd07G15.r OSJNEd Oryza sativa (japonica cultivar-group) cDNA
DEFINITION clone OSJNEd07G15 3', mRNA sequence.
ACCESSION CB662942
VERSION CB662942.1 GI:29666667
KEYWORDS EST.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Eriharoidae; Oryzaceae; Oryza.
1 (bases 1 to 760)
JOURNAL
AUTHORS Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wang,R. and Wang,G.
TITLE Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
JOURNAL
COMMENT Unpublished
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aac cga cgg cca gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 07 row: G column: 15
Seq primer: gga aac agc tat gac cat g.
FEATURES
Source
1..760
location/Qualifiers
1..760
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSJNEd07G15"
/rissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSJNEd"
/note="Vector: pBluescript II KS +; Site.1: EcoRI; Site.2:
XhoI; 24 hrs after inoculation with Rice Blast (CG240-1)"
BASE COUNT 254 a 176 c 158 g 172 t
ORIGIN
Query Match 8.8%, Score 545.8; DB 14; Length 760;

```

```

Best Local Similarity 95.3%; Pred. No. 9e-110;
Matches 573; Conservative 0; Mismatches 27; Indels 1; Gaps 1;
OY 5564 CATATCAATATGTTGATGATGATTTTCAACCAATCCATACAGTCAATCAATCAAGC 5623
Db 1 CATATCAATATGTTGATGATGATTTTCAACCAATCCATACAGTCAATCAATCAAGC 60
OY 5624 AAGTGTATATATATATATATATATATATATATATATATATATATATATATATAT 5683
Db 61 AAGTGTATATATATATATATATATATATATATATATATATATATATATATATATAT 120
OY 5684 GATTTCACCGGTGGGGTGTATATATATATATATATATATATATATATATATATATAT 5743
Db 121 GATTCACCGGGGGGGGTATATATATATATATATATATATATATATATATATATATAT 180
OY 5744 CTAGCTAGCAAAAATATATATATATATATATATATATATATATATATATATATATAT 5803
Db 181 CTAGCTAGCAAAAATATATATATATATATATATATATATATATATATATATATATAT 240
OY 5804 CTCAATATGCGAGGACTTTCATCGTTTCGCGGCAAGAAACAAATCAATGATATG 5863
Db 241 CTCAATATGCGAGGACTTTCATCGTTTCGCGGCAAGAAACAAATCAATGATATG 300
OY 5864 CAGGGAACCAACCAAAAGAGG--CCATCAATCCATCCATCCATCCATCCATCCATCC 5922
Db 301 CAGGGAACCAACCAAAAGAGG--CCATCAATCCATCCATCCATCCATCCATCCATCCAT 360
OY 5922 AAGAGATATATTCACAGCTATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5982
Db 361 AAGAGATATATTCACAGCTATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 420
OY 5983 TCAGGGAAGAAAGAGATCATGATTTGCTGATCTCGGGCGCTGATGTTCTTGGGAG 6042
Db 421 TCAGGGAAGAAAGAGATCATGATTTGCTGATCTCGGGCGCTGATGTTCTTGGGAG 480
OY 6043 TCTTGTGTGAGAGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 6102
Db 481 TCTTGTGTGAGAGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
OY 6103 GCCAGTAAAGAGCAAAAGACAAATATATATATATATATATATATATATATATATAT 6162
Db 541 GCCAGTAAAGAGCAAAAGACAAATATATATATATATATATATATATATATATATAT 600
OY 6163 C 6163
Db 601 C 601

RESULT 7
CB659202 822 bp mRNA linear EST 09-APR-2003
LOCUS OSJNEC15P20.f OSJNEC Oryza sativa (japonica cultivar-group) cDNA
DEFINITION clone OSJNEC15P20 5', mRNA sequence.
ACCESSION CB659202
VERSION CB659202.1 GI:29662927
KEYWORDS EST.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Eriharoidae; Oryzaceae; Oryza.
1 (bases 1 to 822)
JOURNAL
AUTHORS Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wang,R. and Wang,G.
TITLE Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
JOURNAL
COMMENT Unpublished
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967

```



```

source
1. 845
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSJNEC07M21"
/tissue_type="leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_id="OSJNEC"
/note="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2: XhoI; 6 hrs after inoculation with Rice Blast (99240-1)"
BASE COUNT      161 a      243 c      293 g      148 t
ORIGIN
Query Match      8.5%; Score 530; DB 14; Length 845;

```

RESULT 10	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	TITLE	JOURNAL	COMMENT
CB669208		OSJNE0111.f OSJNE <i>Oryza sativa</i> (Japonica cultivar-group) cDNA clone OSJNE0111 5', mRNA sequence.	CB669208		EST.		<i>Oryza sativa</i> (Japonica cultivar-group)	1 (bases 1 to 832)	Kudrna,D., Dean,R., Soderlund,C., Ming,R. and Wang,G.		Arizona Genomics Institute University of Arizona Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ 85721-0088, USA Tel: 520 626 3667 Fax: 520 621 9288 Email: <a href="http://genome.arizona.edu">http://genome.arizona.edu</a> PCR Primers FORWARD: gta aaa cga cgg cca gtc
CB669208		linear	EST	09-APR-2000			<i>Oryza sativa</i> (Japonica cultivar-group)	1 (bases 1 to 832)	Janasuriratt,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur E.		
CB669208		linear	EST	09-APR-2000			<i>Oryza sativa</i> (Japonica cultivar-group)	1 (bases 1 to 832)	Large-scale identification of ESTs involved in the interaction between rice and <i>Magnaporthe oryzae</i>		
CB669208		linear	EST	09-APR-2000			<i>Oryza sativa</i> (Japonica cultivar-group)	1 (bases 1 to 832)	Unpublished		
CB669208		linear	EST	09-APR-2000			<i>Oryza sativa</i> (Japonica cultivar-group)	1 (bases 1 to 832)	Contact: Rod Ming		
CB669208		linear	EST	09-APR-2000			<i>Oryza sativa</i> (Japonica cultivar-group)	1 (bases 1 to 832)	Arizona Genomics Institute		
CB669208		linear	EST	09-APR-2000			<i>Oryza sativa</i> (Japonica cultivar-group)	1 (bases 1 to 832)	University of Arizona		
CB669208		linear	EST	09-APR-2000			<i>Oryza sativa</i> (Japonica cultivar-group)	1 (bases 1 to 832)	Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ		
CB669208		linear	EST	09-APR-2000			<i>Oryza sativa</i> (Japonica cultivar-group)	1 (bases 1 to 832)	85721-0088, USA		
CB669208		linear	EST	09-APR-2000			<i>Oryza sativa</i> (Japonica cultivar-group)	1 (bases 1 to 832)	Tel: 520 626 3667		
CB669208		linear	EST	09-APR-2000			<i>Oryza sativa</i> (Japonica cultivar-group)	1 (bases 1 to 832)	Fax: 520 621 9288		
CB669208		linear	EST	09-APR-2000			<i>Oryza sativa</i> (Japonica cultivar-group)	1 (bases 1 to 832)	Email: <a href="http://genome.arizona.edu">http://genome.arizona.edu</a>		
CB669208		linear	EST	09-APR-2000			<i>Oryza sativa</i> (Japonica cultivar-group)	1 (bases 1 to 832)	PCR Primers		
CB669208		linear	EST	09-APR-2000			<i>Oryza sativa</i> (Japonica cultivar-group)	1 (bases 1 to 832)	FORWARD: gta aaa cga cgg cca gtc		

BACKWARD: gga aac agc tat gac cat g  
Plate: 01 row: 1 column: 11  
Seq primer: gta aac cga cgy cca gty.

FEATURES  
Location/Qualifiers  
1..832

source  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="OSJNB01111"  
/tissue\_type="leaf"  
/dev\_stage="3 week"  
/lab\_host="DH10B"  
/clone\_lib="OSJNB"  
/note="Vector: pBluescript II KS +; Site\_1: EcoRI; Site\_2:  
XhoI; 24 hrs after inoculation with Rice Blast (70-15)"  
BASE COUNT 156 a 237 c 291 g 148 t  
ORIGIN

Query Match 8.3%; Score 516.8; DB 14; Length 832;  
Best Local Similarity 97.5%; Pred. No. 2.5e-103;  
Matches 540; Conservative 0; Mismatches 2; Indels 12; Gaps 1;

Dy 1033 TCCTCTCTCTCTCTCCAGCGCATTCGCGCGAGGTTTCTCCGATCAACCCCTGGAAT 1092  
|||||  
Dy 4 TCCTCTCTCTCTCTCCAGCGCATTCGCGCGAGGTTTCTCCGATCAACCCCTGGAAT 63  
1093 CCCCCCTCGCATTCGATCGAGGAGGTAGCCCGCGCATTCGCGCGAGGAGCGGATTC 1152  
64 CCCCCCTCGCATTCGATCGAGGAGGTAGCCCGCGCATTCGCGCGAGGAGCGGATTC 123  
Dy 1153 CGATTCGCGCATTCGATCGAGGAGGTAGCCCGCGCATTCGCGCGAGGAGCGGATTC 1212  
Dy 124 CGATTCGCGCATTCGATCGAGGAGGTAGCCCGCGCATTCGCGCGAGGAGCGGATTC 183  
Dy 1213 GCTTCGCGCGCATTCGATCGAGGAGGTAGCCCGCGCATTCGCGCGAGGAGCGGATTC 1272  
184 GCTTCGCGCGCATTCGATCGAGGAGGTAGCCCGCGCATTCGCGCGAGGAGCGGATTC 243  
Dy 1273 GAGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGT 1332  
244 GAGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGT 303  
Dy 1333 GTGAGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGT 1392  
304 GTGAGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGT 363  
Dy 1393 CCTAGCCCGAGGAGCGGAGGAGGATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1452  
364 CCTAGCCCGAGGAGCGGAGGAGGATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 423  
Dy 1453 TGTAGAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1500  
424 TGTAGAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 483  
Dy 1501 GAGAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1560  
484 GAGAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 543  
Dy 1561 CGCATGTGAGGAGG 1574  
544 CGCATGTGAGGAGG 557

RESULT 11  
LOCUS BI808598 663 bp mRNA linear EST 02-Oct-2001  
DEFINITION D004E01 Oryza sativa mature leaf library induced by M.grisea Oryza  
sativa cDNA clone D004E01, mRNA sequence.  
ACCESSION BI808598  
VERSION BI808598.1 GI:1565786  
KEYWORDS EST.  
SOURCE Oryza sativa

ORGANISM

Oryza sativa  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE

1 (bases 1 to 663)

AUTHORS

Dong,H.T., Li,D.B., Zhuang,X.F., Dai,C.G., Sun,L.X., Pei,Y.X., Wu  
,H.F., Jiang,Y.X., Yu,F.C., Gao,Q.K. and Lou,Y.C.  
A Gene Expression Screen in Oryza sativa  
Unpublished  
Contact: Dong HT  
Laboratory of Functional Genetics  
Bio-technology Institute of Zhejiang University  
Kaixuan Road 268#, Hangzhou, Zhejiang, P.R.China  
Tel: 0086-571-86892051  
Fax: 0086-571-86961525  
Email: htdong@zjuem.zju.edu.cn  
Seq primer: M13 forward primer.

FEATURES

Location/Qualifiers  
1..663

source  
/organism="Oryza sativa"  
/mol\_type="mRNA"  
/db\_xref="taxon:4530"  
/clone="D004E01"  
/tissue\_type="leaf"  
/dev\_stage="Mature stage"  
/clone\_lib="Oryza sativa mature leaf library induced by  
M.grisea"  
/note="Vector: pSport2"  
BASE COUNT 164 a 174 c 153 g 172 t  
ORIGIN

Query Match 7.4%; Score 459.4; DB 12; Length 663;  
Best Local Similarity 94.5%; Pred. No. 1.2e-90;  
Matches 617; Conservative 0; Mismatches 21; Indels 15; Gaps 13;

Dy 4737 AGGATGAAGGCGTGTTCCTCCGCGCG -TTCTATATGATCATCCCTAGATGCA-TTC 4794  
|||||  
Dy 1 AGGATGAAGGCGTGTTCCTCCGCGCGTTCCTATATGATCATCCCTAGATGCA-TTC 60  
4795 AACAGTCCCATCTGAGAGCAAGTCAAGAGCTGCTGTTCCATCCAAATGA-CCGGA 4853  
61 AACAGTCCCATCTGAGAGCAAGTCAAGAGCTGCTGTTCCATCCAAATGA-CCGGA 119  
Dy 4854 CAATTAATTCGCGTCTGAGAGCAAGTCAAGAGCTGCTGTTCCATCCAAATGA-CCGGA 4912  
120 CAATTAATTCGCGTCTGAGAGCAAGTCAAGAGCTGCTGTTCCATCCAAATGA-CCGGA 178  
Dy 4913 CTTCCTGAGTGAAGAGCAAGTCAAGAGCTGCTGTTCCATCCAAATGA-CCGGA 4972  
179 CTTCCTGAGTGAAGAGCAAGTCAAGAGCTGCTGTTCCATCCAAATGA-CCGGA 237  
Dy 4973 GCCGGCCAGCTTCGCTGAGAGGAGTGGCGAGCTGAGAGCAATCCAGAGAGTGGG 5032  
238 GCC-GCCAGGCTGCTGAGAGGAGTGGCGAGCTGAGAGCAATCCAGAGAGTGGG -C 293  
Dy 5033 GTGGCGGAGTTCGCTGAGAGGAGTGGCGAGCTGAGAGCAATCCAGAGAGTGGG 5092  
294 GTGGCGGAGTTCGCTGAGAGGAGTGGCGAGCTGAGAGCAATCCAGAGAGTGGG 353  
Dy 5093 GAAAGACATCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5152  
354 GAAAGACATCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 411  
Dy 5153 GGTGATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 5212  
412 AGTGTGATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 471  
Dy 5213 AGTACCTCTTCT-TCTAAGGAGTATGATACCTCTGATGATGATGATGATGATGATGATGAT 5271  
472 AGTACCTCTTCT-TCTAAGGAGTATGATACCTCTGATGATGATGATGATGATGATGATGAT 531  
Dy 5272 ATACATTAACCAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 5331

Db 532 ATACACATTAAGCC-AATGTTCAATGAGTGTGATGAGCGCTGCGCCCTAATTGACACATGAT 590

QY 5333 TTCAGGAGAGCTGCTAATGATGATCATTAATTAATTTGATGATGATGCTTCCCTCT 5384  
|||||  
Db 591 TTCAGGAGAGCTGCTAATGATGATCATTAATTAATTTGATGATGATGCTTCCCTCT 643

RESULT 12  
AUI73935 464 bp mRNA linear EST 03-APR-2002  
LOCUS AUI73935  
DEFINITION AUI73935 Rice mature leaf Oryza sativa (japonica cultivar-group)  
CDNA clone S20982, mRNA sequence.  
ACCESSION AUI73935  
VERSION AUI73935.1 GI:13165138  
KEYWORDS EST.  
SOURCE Oryza sativa (japonica cultivar-group)  
ORGANISM Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaceae; Oryza.  
REFERENCE 1 (bases 1 to 464)  
Sasaki, T. and Yamamoto, K.  
Rice cDNA from mature leaf (2001)  
AUTHORS Unpublished  
TITLE National Institute of Agricultural Resources  
JOURNAL Rice Genome Research Program, Kamondai 2-1-2, Tsukuba, Ibaraki  
COMMENT Contact: Takuji Sasaki  
National Institute of Agricultural Resources  
Rice Genome Research Program, Kamondai 2-1-2, Tsukuba, Ibaraki  
305-8602, Japan  
Tel: 81-298-38-7441  
Fax: 81-298-38-7468  
Email: tsasaki@ab.affrc.go.jp, URL: http://rpg.dna.affrc.go.jp/  
PROJECT = "RGP".  
Location/Qualifiers  
1. 464  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="S20982"  
/tissue\_type="mature leaf"  
/clone\_lib="Rice mature leaf"  
BASE COUNT 108 a 104 c 122 g 127 t 3 others  
ORIGIN

Query Match 7.1%; Score 441.2; DB 9; Length 464;  
Best Local Similarity 98.7%; Pred. No. 1.1e-86;  
Matches 443; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4936 GACTTCGTCATGAGCGCTGCTGCGCGGCAAGATGGCGCGGCGGCTGCGGCGG 4995  
|||||  
Db 1 GACTTCGTCATGAGCGCTGCTGCGCGGCAAGATGGCGCGGCGGCTGCGGCGG 60

QY 4996 GTGGGAGAGCTGAGAGCATCTCCAGAGAGATGTGCGTGGCGGCGGCTTCTGCTGGCTCG 5055  
|||||  
Db 61 GTGGGAGAGCTGAGAGCATCTCCAGAGAGATGTGCGTGGCGGCGGCTTCTGCTGGCTCG 120

QY 5056 ACCTCTGAGACCGAACCAGACAGCTGCTTGGCTTGAAGAAGACCTGGAGAGGAAAT 5115  
|||||  
Db 121 ACCTCTGAGACCGAACCAGACAGCTGCTTGGCTTGAAGAAGACCTGGAGAGGAAAT 180

QY 5116 GATGATTGAGACAGCGCTAGACATTTCTATGCTGTGAGGTGATGATTGTGGTGCACGT 5175  
|||||  
Db 181 GATGATTGAGACAGCGCTAGACATTTCTATGCTGTGAGGTGATGATTGTGGTGCACGT 240

QY 5176 GGTGTCATTTGTCATTTCTTTGTAAGGAGACCTTTAGTACCTCTCTCTTAAGGAGAC 5235  
|||||  
Db 241 GGTGTCATTTGTCATTTCTTTGTAAGGAGACCTTTAGTACCTCTCTCTTAAGGAGAC 300

QY 5236 TTATGACCCCTGTTGATGATCTCATGCTGCTAATTAATTAATTAATTAATTAATTAAT 5295  
|||||  
Db 301 TTATGACCCCTGTTGATGATCTCATGCTGCTAATTAATTAATTAATTAATTAATTAAT 360

QY 5296 GATGATGAGCGTCTGCTGCTAATTTGAAGAGACTGATTTGAGGAGCTGCTGATGATCAT 5355

Db 361 GGTGATGAGCGTGCCTGCTAATTTGAACAGATGATTCAGGACACTGATGCTATCATCAT 420  
|||||

QY 5356 TCAATTAATTTGATGATGATGCTTCCCTCT 5384  
|||||  
Db 421 TCAATTAATTTGATGATGATGCTTCCCTCT 449

RESULT 13  
BI809493 566 bp mRNA linear EST 02-OCT-2001  
LOCUS BI809493  
DEFINITION F010F11 Oryza sativa mature leaf library induced by M.grisea Oryza  
sativa cDNA clone F010F11, mRNA sequence.  
ACCESSION BI809493  
VERSION BI809493.1 GI:15856681  
KEYWORDS EST.  
SOURCE Oryza sativa  
ORGANISM Oryza sativa  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaceae; Oryza.  
REFERENCE 1 (bases 1 to 566)  
Dong, H.T., Li, D.B., Zhuang, X.F., Dai, C.G., Sun, L.X., Pei, Y.X., Wu  
H.F., Jiang, Y.X., Yu, F.C., Gao, Q.K. and Lou, Y.C.  
A Gene Expression Screen in Oryza sativa  
AUTHORS Unpublished  
TITLE Laboratory of Functional Genetics  
JOURNAL Bio-technology Institute of Zhejiang University  
COMMENT Contact: Dong HT  
Laboratory of Functional Genetics  
Bio-technology Institute of Zhejiang University  
Kaixuan Road 268#, Hangzhou, Zhejiang, P.R.China  
Tel: 0086-571-86892051  
Fax: 0086-571-86961525  
Email: hcdong@zjuem.zju.edu.cn  
Seq primer: M13 forward primer.  
Location/Qualifiers  
1. 566  
/organism="Oryza sativa"  
/mol\_type="mRNA"  
/db\_xref="taxon:4530"  
/clone="F010F11"  
/tissue\_type="leaf"  
/dev\_stage="Mature stage"  
/clone\_lib="Oryza sativa mature leaf library induced by  
M.grisea"  
/note="Vector: pSport2"  
BASE COUNT 145 a 150 c 134 g 137 t  
ORIGIN

Query Match 6.8%; Score 422.8; DB 12; Length 566;  
Best Local Similarity 97.7%; Pred. No. 1.5e-82;  
Matches 502; Conservative 0; Mismatches 4; Indels 8; Gaps 7;

QY 4829 CTGTTCCTCATCCAGATGACCCGAGACATTAATCTTCTGCTACTACACAGCATCGGAGTA 4888  
|||||  
Db 39 CTGTTCCTCATCCAGATGACCCGAGACATTAATCTTCTGCTACTACACAGCATCGGAGTA 98

QY 4889 ACACACTCATGATGCCGACATTAATCTTCTGCTACTACACAGCATCGGAGTA 4948  
|||||  
Db 99 ACACACTCATGATGCCGACATTAATCTTCTGCTACTACACAGCATCGGAGTA 158

QY 4949 GGGCTGCGGCTGGCGGCAAGATTTGGCCGCGGAGCTGCTGTCAGCGGGGAGAGCTGG 5008  
|||||  
Db 159 GGGCTGCGGCTGGCGGCAAGATTTGGCCGCGGAGCTGCTGTCAGCGGGGAGAGCTGG 216

QY 5009 AGCATCTCCAGAGAGAGATGTGCGGTGGCGGCTTCTGCTGCTGCTGCTGAGAC 5068  
|||||  
Db 217 AGCATCTCCAGAGAGAGATGTGCGGTGGCGGCTTCTGCTGCTGCTGCTGAGAC 274

QY 5069 GAAACCCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5127  
|||||  
Db 275 GAAACCCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 334

QY 5128 GAGCGAGACA-TTGCTAATGCTGTGAGGTGATGATGTTGGTGTGCTGCTGCTGCTA-TT 5185



|||||  
 Db 335 GGCATGACACTTGCATAGCTGAGGTGATGATGTGGTGTGCTGTGCTGTGACTT 394  
 |||||  
 QY 5186 GTGATTCCTTTGTAAGGACACCTCTTAGTACCTCTTCTTAAAGGACTTAGTACCC 5245  
 |||||  
 Db 395 GTGATTCCTTTGTAAGGACACCTCTTAGTACCTCTTCTTAAAGGACTTAGTACCC 454  
 |||||  
 QY 5246 TTGTGATTCATGCTGCTTAATATATATACATTAAGCAAAATGTCATTGATGATGG 5305  
 |||||  
 Db 455 TTGTGATTCATGCTGCTTAATATATATACATTAAGCAAAATGTCATTGATGATGG 513  
 |||||  
 QY 5306 GGTGCTCCTTAATTTGACGACTGATTTGAGGCA 5339  
 |||||  
 Db 514 GGTGCTCCTTAATTTGACGACTGATTTGAGGCA 547  
 |||||

## RESULT 14

BQ908360 421 bp mRNA linear EST 19-AUG-2002  
 LOCUS T006G11 Oryza sativa mature leaf library induced by M.grisea Oryza  
 DEFINITION sativa cDNA clone T006G11, mRNA sequence.  
 ACCESSION BQ908360  
 VERSION BQ908360.1 GI:22307138  
 KEYWORDS EST  
 SOURCE Oryza sativa  
 ORGANISM Oryza sativa  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.

## REFERENCE

1 (bases 1 to 421)  
 Dong,H.T., Li,P.B., Zhuang,X.F., Dai,C.G., Sun,L.X., Pei,Y.X., Wu  
 H.F., Jiang,Y.X., Yu,F.C., Gao,Q.K. and Lou,Y.C.  
 A Gene Expression Screen in Oryza sativa

## TITLE

Unpublished

## JOURNAL

Unpublished

## COMMENT

Contact: Dong HT  
 Laboratory of Functional Genetics  
 Bio-technology Institute of Zhejiang University  
 Kaixuan Road 2668, Hangzhou, Zhejiang, P.R.China  
 Tel: 0086-571-86892051  
 Fax: 0086-571-86961525  
 Email: htdong@zjuem.zju.edu.cn  
 Seq primer: M13 forward primer.  
 Location/Qualifiers

## FEATURES

## source

1..421  
 /organism="Oryza sativa"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4530"  
 /clone="T006G11"  
 /tissue\_type="leaf"  
 /dev\_stage="Mature stage"  
 /clone\_lib="Oryza sativa mature leaf library induced by  
 M.grisea"  
 /note="Vector: pSport2"

BASE COUNT 103 a 93 c 110 g 115 t  
 ORIGIN

Query Match 6.5%; Score 407.4; DB 13; Length 421;  
 Best Local Similarity 99.8%; Pred. No. 3.4e-79;  
 Matches 408; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

|||||  
 Db 181 GCGTAGGTGATGATGATGTTGTGCTGTGCTGTGCTATGTCATTCTTTAGAGGAC 240  
 |||||  
 QY 5206 ACTCTTAGTACCTCTTCTTCTTAAAGGACTTAGTACCTCTTGTGATCTATGCTCTA 5265  
 |||||  
 Db 241 ACTCTTAGTACCTCTTCTTCTTAAAGGACTTAGTACCTCTTGTGATCTATGCTCTA 300  
 |||||  
 QY 5266 AATACATACATACATACGCAAAATGTCATTGATGATGGGCGTCGCCATTAATGAGC 5325  
 |||||  
 Db 301 AATACATACATACATACGCAAAATGTCATTGATGATGGGCGTCGCCATTAATGAGC 360  
 |||||  
 QY 5326 ACTGATTGAGGAGCTGCTAGTATGCTATCATTTAATTAATTTGATGCA 5374  
 |||||  
 Db 361 ACTGATTGAGGAGCTGCTAGTATGCTATCATTTAATTAATTTGATGCA 409  
 |||||

## RESULT 15

CA757602 813 bp mRNA linear EST 27-NOV-2002  
 LOCUS OE04A06-T3.ab1 OE Oryza sativa cDNA clone OE04A06-T3.ab1 similar to  
 DEFINITION No protein alignment, mRNA sequence.  
 ACCESSION CA757602  
 VERSION CA757602.1 GI:25801641  
 KEYWORDS EST  
 SOURCE Oryza sativa  
 ORGANISM Oryza sativa  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.

## REFERENCE

1 (bases 1 to 813)  
 Bolnert,H.J., Borchert,C., Brazille,S., Brooks,J., Eaton,M., Ferreira  
 H., Kawasaki,S., McColough,A., Michalowski,C.B., Palacios,C.,  
 Seara,G., Wheeler,M. and Zepeda,G.R.  
 Functional Genomics of Plant Stress Tolerance

## TITLE

Unpublished

## JOURNAL

Unpublished

## COMMENT

Contact: Mark Fredrickson  
 Department of Plant Biology  
 University of Illinois  
 1201 W. Gregory Dr., Urbana, IL 61801, USA  
 Tel: 2172653473  
 Email: bolnertlab@life.uiuc.edu.  
 Location/Qualifiers

## FEATURES

## source

1..813  
 /organism="Oryza sativa"  
 /mol\_type="mRNA"  
 /strain="pokali"  
 /db\_xref="taxon:4530"  
 /clone="OE04A06-T3.ab1"  
 /tissue\_type="roots"  
 /dev\_stage="1 week"  
 /clone\_lib="OE"  
 /note="2-3 d 150mM NaCl"

BASE COUNT 183 a 177 c 206 g 246 t 1 others  
 ORIGIN

Query Match 5.7%; Score 354; DB 14; Length 813;  
 Best Local Similarity 90.1%; Pred. No. 3e-67;  
 Matches 401; Conservative 0; Mismatches 41; Indels 3; Gaps 2;

|||||  
 Db 5774 AATCAATTCACATCCACGCGGATGGAACAAAGACAGATGTCACGCTATCTTCG 5831  
 |||||  
 QY 5832 GCGGCAAGAAAGATCAAGTATATACATTGCGAGGAAACACAAAGAAAG-CCATCC 5890  
 |||||  
 Db 683 GCTGCGCATCAACGATCCAGTATACATTGCGAGGGAACCCCAAGAGGCCATCC 624  
 |||||  
 QY 5891 AATCAATTCACATCCACGCGGATGGAACAAAGACAGATGTCACGCTATCTTCG 5950  
 |||||

```

Db      623 AATCCACTCACTCCACCGGCATGAGACAGACAGATGCTCACAGCTATCTTCTG 564
QY      5951 CTTCTACAGATTGATACCTTGTACTGTCTTCAGGCAAAAAGAGCATCAGATTAGTC 6010
      |||||||
Db      563 CTTCTACAGATTGATACCTTGTACTGTCTTCAGGCAAAAAGAGCATCAGATTAGTC 504
QY      6011 TGAATCGGCGCGCTGAGTCTTGTGAGAGATCTTGTGTGAGTGCAGAGTGAAGA 6070
      |||||||
Db      503 TGAATCGGCGCGCTGAGTCTTGTGAGAGATCTTGTGTGAGTGCAGAGTGAAGA 444
QY      6071 TCGGCTGCGCGCTTCTCTACCGAATCGCGAGTAAGAGCCAAAAGACAATAA 6130
      |||||||
Db      443 TCGGCTGCGCGCTTCTCTACCGAATCGCGAGTAAGAGCCAAAAGACAATAA 384
QY      6131 TACGCAATGCGGATCGCCCATCTG 6155
      |||||||
Db      383 TACGCAATGCGGATCGCCCATCTG 359

```

Search completed: September 14, 2003, 02:18:09  
 Job time : 8204.38 secs



```

; PRIOR FILING DATE: 2000-11-27
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 15418
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-763-203-I

Query Match          1.2%; Score 72; DB 4; Length 15418;
Best Local Similarity   58.6%; Pred. No. 1,6e+08;
Matches    143; Conservative      0; Mismatches     100; Indels       1; Gaps        1.

QY         ATATAGTATGACATACTAATTACGTAAATACCGTAATACCAATGAAGAATTTAACCATAT 461
            ||| |||| | |||| | |||| | ||| |||| | ||| |||| | |||
Db         ATTACTTATATATGGTATATTTCGTATATCCTGTATATTATATATATATATGCAT 1137

QY         AAGGATATACATATATATATATATATATATATATATATATATATATATATATATG 521
            ||| |||| | |||| | |||| | |||| | |||| | |||| | |||
Db         ATA-TATTTATATATGATATATATATATATATATATATATATATATATATATATAT 1072

QY         GTGGCACATGGTAAAAAGTTTTATATATATGAACGATGAGTGACATCCACAAGGCCOA 581
            ||| |||| | |||| | |||| | ||| | ||| | ||| | |||
Db         TATATTATATATATATATATATATATATATATATATATATATATATATATATATAT 1011

QY         TAGTTTTCTGGCACTGTGAGTAGATACGAATGCACCATTATATATCCAATAAATGAMATTAT 641
            ||| |||| | |||| | |||| | |||| | |||| | |||| | |||
Db         TTATATATATATATATATATATATATATATATATATATATATATATATATATATAT 952

QY         GAATAT 645
            |||| 
Db         951 MAT 948

RESULT 3
Sequence 1, Application US/08965729A           |
Patent No. 620751                             |

GENERAL INFORMATION:
APPLICANT: Jian-Ming Gu and Charles T. Esmun
TITLE OF INVENTION: ENDOTHELIAL SPECIFIC EXPRESSION
TITLE OF INVENTION: REGULATED BY EPIC CONTROL ELEMENTS
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center, 1201 West Peachtree Street
CITY: Atlanta
STATE: GA
COUNTRY: USA
ZIP: 30309-4530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0., Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/965,729A
FILING DATE: 07-NOV-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/030,718
FILING DATE: 08-NOV-1997
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: OMRP 164 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-873-8794
TELEFAX: 404-873-8795
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3481 base pairs
TYPE: nucleic acid
```

```

STRANDEDNESS: single
TOPOLOGY: Linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
  NAME/KEY: misc.feature
  OTHER INFORMATION: /note= "Nucleotides 2270 through 2840 are a
  OTHER INFORMATION: large endothelial specific element"; murine
  FEATURE:
  NAME/KEY: misc.feature
  OTHER INFORMATION: /note= "Nucleotides 2990 through 3061 are a
  OTHER INFORMATION: serum response element"; murine
  FEATURE:
  NAME/KEY: misc.feature
  OTHER INFORMATION: /note= "Nucleotides 3007 through 3014 are a
  OTHER INFORMATION: thrombin responsive element"; murine
  FEATURE:
  NAME/KEY: misc.feature
  OTHER INFORMATION: /note= "Nucleotides 3130 through 3150 are an
  OTHER INFORMATION: endothelial specific element"; murine
US-08-965-723A-1

Query Match
Best Local Similarity 1.1%; Score 66.8; DB 3; Length 3481;
Matches 94; Conservative 0; Mismatches 27; Indels 1; Gaps 1

QY 398 TTCTCATATGATATATACATATATATACGATATATACGTAATACATATACGTAATAC
  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db 2041 TATCTAGACACAGATATATATATATATATATATATATATATATATATATATATATAC
  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

QY 458 AATATGCGTATATACATATATATATATATATATATATATATATATATATATATATATAT
  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db 2101 GATATATGATATATATATATATATATATATATATATATATATATATATATATATATAT
  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

QY 518 TG 519
  1 1
Db 2160 TG 2161

RESULT 4
US-09-647-826-1/c
Sequence 1, Application US/09647826
Patent No. 6437216
GENERAL INFORMATION:
  APPLICANT: Duff, Gordon W.
  APPLICANT: Nicklin, Martin
  TITLE OF INVENTION: TRANSGENIC MODELS OF INFLAMMATORY DISEASE
  FILE REFERENCE: MSA-003.01
  CURRENT APPLICATION NUMBER: US/09/647,826
  CURRENT FILING DATE: 2001-03-12
  PRIOR APPLICATION NUMBER: GB 9723835.6
  PRIOR FILING DATE: 1997-11-13
  NUMBER OF SEQ ID NOS: 2
  SOFTWARE: PatentIn Ver. 2.0
  SEQ ID NO 1
  LENGTH: 6350
  TYPE: DNA
  ORGANISM: mouse
US-09-647-826-1

Query Match
Best Local Similarity 1.1%; Score 66.2; DB 4; Length 6350;
Matches 95; Conservative 0; Mismatches 48; Indels 0; Gaps 0

QY 368 TCTAATTCACACGAGCTACTCCCTCGTTGATATATGTAATATACATATATACGTAT
  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db 784 TGTCTTGAATCAAGATATACACACACACACACACACACACACACATATATATATATAT
  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

QY 428 AATATGCGTATATACATATATATATATATATATATATATATATATATATATATATATAT
  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db 724 AATATGCGTATATATATATATATATATATATATATATATATATATATATATATATATAC
  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

724 AATATGCGTATATATATATATATATATATATATATATATATATATATATATATATATAC 665

```

```

OY      488 TATATAATTAATTGCTGGTGT 510
        ||| ||| ||| ||| |||| |
Db       664 CATACATCATCAATATGTGART 642


RESULT 5
US-09-783-203-1
; Sequence 1, Application US/09783203
; Patent No. 6576464
; GENERAL INFORMATION:
; APPLICANT: Genon Corporation
; APPLICANT: Gold, Joseph
; APPLICANT: Lebkowski, Jane
; TITLE OF INVENTION: Tracked stem cells
; FILE REFERENCE: 096/003
; CURRENT APPLICATION NUMBER: US/09/783,203
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/253,443
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/253,357
; PRIOR FILING DATE: 2000-11-27
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 1
; LENGTH: 15418
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-783-203-1

Query Match          1.1%; Score 65.8; DB 4; Length 15418;
Best Local Similarity 70.3%; Pred. No. 7,9e-07;
Matches 102; Conservative 0; Mismatches 42; Indels 1; Gaps 1;

OY      402 ATAAATGAATATCAGATATATACGTTATATACGTATATACACAATACGTTATACATAT 461
        ||| ||| ||| |||| |||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Dd       1031 ACAATATGAAAATGCATATATTAATTAATATACATATATTAATATTCATATATTAATATATAT 1090

OY      462 AATCG-TAATACACTAATATATATATATATATATATATATATATATATATATNGTGCTGTANGSG 520
        ||| ||| |||| |||||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||
Dd       1091 ATAATATATACATATATATATATATATATATATATATATATATATATATATATATATAT 1150

OY      521 GGTCGCAATGCTAAAAAGTTTTATA 545
        ||| ||| ||| ||| |||| 
Dd       1151 ATTACAGAATATATACAAATATATTA 1175


RESULT 6
US-08-947-823-1
; Sequence 1, Application US/08947823
; Patent No. 6114605
; GENERAL INFORMATION:
; APPLICANT: Williamson, Valerie M.
; APPLICANT: Kaloshian, Isqouhi
; APPLICANT: Yaghoobi, Jafar
; APPLICANT: Bodeau, John
; APPLICANT: Milligan, Stephen
; TITLE OF INVENTION: Procedures and Materials for Confering
; TITLE OF INVENTION: Pest Resistance in Plants
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
```

```

APPLICATION NUMBER: US/08/947,823
FILING DATE: 09-OCT-1997
CLASSIFICATION: 800
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/18802
FILING DATE: 09-OCT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/028,191
FILING DATE: 10-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Bastian, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 023070-070210US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEX: (415) 576-0300
INFORMATION FOR SEQ ID NO.: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 51953 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-947-823-1

Query Match      1.1%; Score 65.8; DB 3; Length 51952;
Best local similarity 51.3%; Pred. No. 1.8e-06;
Matches 177; Conservative 0; Mismatches 167; Indels 1; Gaps 1;

QY    408  GTATATACATATATATAGTATATACGTATATACGATATATACATATATATATGGTA 467
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB    13653 GCATCCCGAGAAAACCTGTGTATAATATATATATATATATATATATATAT-ATA 13711

QY    468  TATCATATATATATATATATATATATATATATATATATATATATATATATATATGTCGA 527
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB    13712 TATATATATATATATATATATATATATATATATATATATATATATTTCTATTAAAAAGAATGC 13777

QY    528  ATGCTAATAAAAGTTTTATATATATGAAAGCATGAGTAGTACTATCCACTATGACCCTATAGTT 587
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB    13772 CTATTTTCATTTTATATATGCTTTAAAAAAGAAATTAATCTCTTTCTTTTGACAATVATT 13833

QY    588  TCTGCAACGTGTAGATATATACGAATGACACAAATTATATCCATATAAATTGATATATATATTC 647
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB    13832 TAACCTTAACTTTCCACGTATACATGTTTATAGACAAACAAATATTAATGACATTTTAACTTT 13899

QY    648  CTCGCGACGAAAAATTAAGACATATATATGCGATATACCATTTATCCAGATATATCTAAAT 707
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB    13892 GTACATATAGAAAAGTAAACATATGATATGCTGTGCCCTAAACATGATAGATGTATAA 13951

QY    708  TTCACGTATATATCTAAATTCOCACCTGATCCCTTTTATGATATA 752
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB    13952 TTCAAAGTCATGAAATGATTTTATGATATATATATTAAGATATA 13996

DESCRIPT 7
US-08-866-340-1/c
Sequence 1, Application US/08866540
Patent No. 6020518
GENERAL INFORMATION:
APPLICANT: Szyf, Moshe
APPLICANT: Bigey, Pascal
APPLICANT: Ramchandani, Shyam
TITLE OF INVENTION: DNA METHYLTRANSFERASE GENOMIC
TITLE OF INVENTION: SEQUENCES AND ANTISENSE OLIGONUCLEOTIDES
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: HALE AND DORR LLP
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: United States of America
ZIP: 02109
COMPUTER READABLE FORM:

```

Query Match	1.0%;	Score 55;	DB 3;	Length 4400;
Best Local Similarity	71.1%;	Pred. No. 5,7e-07;		
Matches	86;	Conservative	0;	Mismatches 35; Indels 0; Gaps 0

402 AATATGATATACATATACGTTATATACGATATACATATACGATATATACATAT 461

Query Match 1.0%; Score 64.8; DB 3; Length 80595;  
 Best Local Similarity 70.3%; Pred. No. 4.5e-06;  
 Matches 101; Conservative 0; Mismatches 42; Indels 1; Gaps 1

QY 402 ATATATGTATATACATATACGTATATATACGTATATACATATACATATACATAT 461  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 15548 ATATATGTATATACATATACGTATATATACATATATATATATATATATAT 15607  
462 ATGTATATACAT 521  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 15608 GT-GTATATATACATATATATATATATATATATATATATATATATATATAT 15666  
QY 522 GTGGCAATGCTAAAGTTTATA 545  
- ||||| |||||  
DB 15667 TATATAGACAT 15690

## RESULT 11

US-09-078-294-4/c  
; Sequence 4, Application US/09078294  
; Patent No. 6265211  
; GENERAL INFORMATION:  
; APPLICANT: Choo, Kong-Hong Andy  
; APPLICANT: Du Sart, Desiree  
; APPLICANT: Cancilla, Michael R.  
; TITLE OF INVENTION: A NOVEL NUCLEIC ACID MOLECULE  
; FILE REFERENCE: Davies Col  
; CURRENT APPLICATION NUMBER: US/09/078,294  
; CURRENT FILING DATE: 1998-05-13  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 80246  
; TYPE: DNA  
; ORGANISM: Nucleotide sequence of NC-contig  
US-09-078-294-4

Query Match 1.0%; Score 64.6; DB 3; Length 80246;  
Best Local Similarity 69.4%; Pred. No. 5.1e-06;  
Matches 102; Conservative 0; Mismatches 44; Indels 1; Gaps 1;

QY 399 TTCAATATGTATATACATATATACGTATATATATACGTATATACATATACGTATATAC 458  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 15415 TTATGTCTATATACATATATATATATATATATATATATATATATATATATAT 15356  
459 TATATGTATATACAT 518  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 15355 TGTCTA-TATACACATATATACATATATATATATATATATATATATATATAT 15297  
QY 519 GGGGTGGCAATGCTAAAGTTTATA 545  
- ||||| |||||  
DB 15296 GTACATACATATATATACATATATATATATATATATATATATATATATAT 15270

## RESULT 12

US-09-078-294-3/c  
; Sequence 3, Application US/09078294  
; Patent No. 6265211  
; GENERAL INFORMATION:  
; APPLICANT: Choo, Kong-Hong Andy  
; APPLICANT: Du Sart, Desiree  
; APPLICANT: Cancilla, Michael R.  
; TITLE OF INVENTION: A NOVEL NUCLEIC ACID MOLECULE  
; FILE REFERENCE: Davies Col  
; CURRENT APPLICATION NUMBER: US/09/078,294  
; CURRENT FILING DATE: 1998-05-13  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 80595  
; TYPE: DNA  
; ORGANISM: Nucleotide sequence of HC-contig  
US-09-078-294-3

Query Match 1.0%; Score 64.6; DB 3; Length 80595;  
Best Local Similarity 69.4%; Pred. No. 5.1e-06;  
Matches 102; Conservative 0; Mismatches 44; Indels 1; Gaps 1;

QY 399 TTCAATATGTATATACATATATACGTATATATATACGTATATATACATATACGTATATAC 458  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 15680 TTATGTCTATATACATATATATATATATATATATATATATATATATATATAT 15621  
QY 459 TATATGTATATACAT 518  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 15620 TGTCTA-TATACACATATATACATATATATATATATATATATATATATATAT 15562  
QY 519 GGGGTGGCAATGCTAAAGTTTATA 545  
- ||||| |||||  
DB 15561 GTACATACATATATATACATATATATATATATATATATATATATATATAT 15535

## RESULT 13

US-09-461-325-77/c  
; Sequence 77, Application US/09461325A  
; Patent No. 6475753  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: 94 Human Secreted Proteins  
; FILE REFERENCE: P2029P1  
; CURRENT APPLICATION NUMBER: US/09/461,325A  
; CURRENT FILING DATE: 1999-12-14  
; EARLIER APPLICATION NUMBER: PCT/US99/13418  
; EARLIER FILING DATE: 1999-06-15  
; EARLIER APPLICATION NUMBER: 60/089,507  
; EARLIER FILING DATE: 1998-06-16  
; EARLIER APPLICATION NUMBER: 60/089,508  
; EARLIER FILING DATE: 1998-06-16  
; EARLIER APPLICATION NUMBER: 60/089,509  
; EARLIER FILING DATE: 1998-06-16  
; EARLIER APPLICATION NUMBER: 60/089,510  
; EARLIER FILING DATE: 1998-06-16  
; EARLIER APPLICATION NUMBER: 60/090,112  
; EARLIER FILING DATE: 1998-06-22  
; EARLIER APPLICATION NUMBER: 60/090,113  
; EARLIER FILING DATE: 1998-06-22  
; NUMBER OF SEQ ID NOS: 332  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 77  
; LENGTH: 1592  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-461-325-77

Query Match 1.0%; Score 64.4; DB 4; Length 1592;  
Best Local Similarity 77.1%; Pred. No. 4.1e-07;  
Matches 91; Conservative 0; Mismatches 26; Indels 1; Gaps 1;

QY 402 ATATATGTATATACATATATACGTATATATATACGTATATATACATATATACGTATATAC 461  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 539 ATATGTGTGTATATACATATATATATATATATATATATATATATATATATATAT 480  
QY 462 ATGTATATACAT 519  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 479 AT-GTGTCTAT 423

## RESULT 14

US-08-232-463-14  
; Sequence 14, Application US/08232463  
; Patent No. 5670367  
; GENERAL INFORMATION:  
; APPLICANT: DORNER, F.  
; APPLICANT: SCHRIFFLINGER, F.  
; APPLICANT: FALKNER, F. G.  
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 1800 Diagonal Road, Suite 500  
; CITY: Alexandria

```

STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PT9PL-F15
US-08-232-463-14

Query Match
Best Local Similarity 6.6%; Score 63.6; DB 1; Length 7218;
Matches 27; Conservative 222; Mismatches 161; Indels 0; Gaps 0;

QY 2309 CCTATTATGTCCTCCAGAAATTGATGATCCCGCCCTTCAATTGCTGAATTAAGCTTA 2368
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1074 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1133
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2369 TTTGTTGTCGCCCTCACTTTTCCAGCTTATGTTGTCGACATGTCGACATGAGCCT 2428
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1134 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1193
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2429 GCTTTGTCGCCCTAATTTAATTTGCAATTCATTTTGACATGATTAATAAGACA 2488
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1194 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1253
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2489 CTAGAGCAACATGCACATGCTGCTATNCCATATACATTTAGTGTCTGTCATTTTA 2548
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1254 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1313
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2549 TCATGTCACCTCATGTCATTTTCATCTCAGGCTCTGACACATCCACCTGAGGCTC 2608
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1314 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1373
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2609 TGGAGCTGCTCATTCTTCTGTAATGACAGATGTTACACAGAGCGCTCCTAATTCAT 2668
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1374 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1433
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2669 CGGAGGTACTTACTTACTGTTGTTACATTTTCAGATTTGTATGAACATAC 2718
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1434 YGTACCAATTCCTTCTATCTTTTAACTGACATAGATAGTAAATAC 1483
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 15  
 US-09-671-317-141/c  
 ; Sequence 141, Application US/09671317  
 ; Patent No. 6528260  
 ; GENERAL INFORMATION:

```

APPLICANT: Blumenfeld, Marta
APPLICANT: Chumakov, Ilya
APPLICANT: Bouqueleret, Lydie
APPLICANT: Cohen, Amick
TITLE OF INVENTION: BIALLELIC MARKERS RELATED TO GENES INVOLVED IN DRUG METABOLISM
FILE REFERENCE: 62,053.CIP
CURRENT APPLICATION NUMBER: US/09/671,317
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US 09/536,178
PRIOR FILING DATE: 2000-03-23
PRIOR APPLICATION NUMBER: PCT/IB00/00403
PRIOR FILING DATE: 2000-03-24
PRIOR APPLICATION NUMBER: US 60/126,269
PRIOR FILING DATE: 1999-03-25
PRIOR APPLICATION NUMBER: US 60/131,961
PRIOR FILING DATE: 1999-04-30
NUMBER OF SEQ ID NOS: 977
SOFTWARE: Patent.pm
SEQ ID NO 141
LENGTH: 1001
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: allele
LOCATION: 501
OTHER INFORMATION: 10-436-43 : polymorphic base G or C
NAME/KEY: misc_binding
LOCATION: 482..500
OTHER INFORMATION: 10-436-43.mis1
NAME/KEY: misc_binding
LOCATION: 502..521
OTHER INFORMATION: 10-436-43.mis2, potential complement
NAME/KEY: primer_bind
LOCATION: 459..476
OTHER INFORMATION: upstream amplification primer
NAME/KEY: primer_bind
LOCATION: 859..878
OTHER INFORMATION: downstream amplification primer, complement
NAME/KEY: misc_binding
LOCATION: 489..513
OTHER INFORMATION: 10-436-43 potential probe
NAME/KEY: misc_feature
LOCATION: 297,539,629,650,976..1001
OTHER INFORMATION: n=a, g, c or t
US-09-671-317-141

```

```

Query Match
Best Local Similarity 78.4%; Score 63.4; DB 4; Length 1001;
Matches 76; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 408 GTATATACATATATACGATATATATACATATACATATATACATATATACGTA 467
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 222 GTATATACGTCGTGTGTATATATATATATATACACACACATATATATATATA 163
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 468 TATACATATATATATATATATATATATATATATATATATATATATATATAT 504
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 162 TATACGATATATATATATATATATATATATATATATATATATATATATAT 126
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Search completed: September 14, 2003, 02:23:41  
 Job time : 281.197 secs